



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Présentée et soutenue par :

Aurore Sanchez

Le jeudi 12 juin 2014

Titre :

La ségrégation du plasmide F d'Escherichia coli : Etude des spécificités d'interaction du centromère avec la protéine SopB et organisation du complexe de partition étendu

ED BSB : Microbiologie

Unité de recherche :

Laboratoire de Microbiologie et Génétique Moléculaires

Directeur(s) de Thèse :

M. Jean-Yves Bouet, Chargé de recherche, CNRS, Toulouse, France

Rapporteurs :

M. Hallet Bernard, Professeur UCL, Louvain, Belgique

M. François-Xavier Barre, Directeur de recherche, CNRS, Gif-sur-Yvette, France

Autre(s) membre(s) du jury :

M. Gutierrez Claude, Professeur UT3 Paul Sabatier, Toulouse, France

M. Parmeggiani Andrea, Professeur Université Montpellier II, Montpellier, France

Résumé

La ségrégation du matériel génétique est une étape fondamentale du cycle cellulaire permettant la transmission du patrimoine génétique au cours des générations. Dans les cellules eucaryotes, la mitose est l'étape qui permet la répartition des chromosomes dupliqués dans chaque cellule fille. Des systèmes actifs, dédiés à la ségrégation de l'ADN sont retrouvés sur la majorité des plasmides et chromosomes bactériens. Ces systèmes ParABS, dits "de partition", sont constitués de deux protéines, ParA et ParB, et d'une séquence centromérique, *parS*. La protéine ParA est une ATPase capable de positionner les plasmides dans le cytoplasme tout au long du cycle cellulaire. Son comportement dynamique fait d'elle le moteur de la partition. La protéine ParB, l'autre acteur de la partition est une protéine adaptatrice entre la molécule d'ADN et la protéine motrice. ParB se fixe sur le centromère *parS* pour former un complexe nucléoprotéique appelée "complexe de partition". En utilisant différents modes de fixation à l'ADN et en établissant des interactions protéine-protéine multiples, ParB est aussi capable de s'organiser en complexe de partition dit "étendu" qui est le substrat de la réaction de ségrégation. La formation du complexe de partition "étendu" est la première étape du mécanisme de partition et est essentielle au processus de ségrégation. L'architecture de ce complexe n'est connue pour aucun des systèmes de partition parABS.

L'un des systèmes modèles majeurs du mécanisme de ségrégation de l'ADN chez les bactéries est le système sopABC du plasmide F d'*E. coli*. Ainsi, au cours de ma thèse, j'ai initié plusieurs projets en parallèle, visant à caractériser à la fois *in vivo* et *in vitro*, les différentes interactions impliquées dans l'organisation du complexe de partition et du complexe de partition "étendu" de ce plasmide. En collaboration avec l'équipe du Dr Véronique Le Berre au Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés (LISBP-INSA) de Toulouse, j'ai contribué à la validation des déterminants nucléiques impliqués dans l'interaction de notre séquence centromérique modèle, *parS*, avec le motif hélice-tour-hélice (HTH) de ParB. Ensuite, nous avons identifié un nouveau motif en dehors du motif HTH et montré qu'une arginine de ce motif est essentielle à l'interaction spécifique avec le centromère. Ces résultats ont révélé une caractéristique conservée dans le règne bactérien : les protéines ParB contiennent un domaine de liaison au centromère, composé de deux motifs séparés et essentiels. Le cœur de mes travaux de thèse a été de comprendre l'organisation du complexe de partition "étendu" et son rôle dans la partition. Par une combinaison d'approche moléculaire *in vivo* et *in vitro*, j'ai

montré que l'architecture de ce complexe étendu en dehors des sites *parS*, nécessitait deux types d'interaction, des interactions protéine/protéine mais également ADN/protéine. Afin d'étudier le mode d'interaction de ParB avec des séquences nucléiques non spécifiques avoisinant le complexe de partition, j'ai mis en œuvre une technique d'immunoprécipitation de chromatine, couplée à des techniques de hautes détections (qPCR ou ChIPseq). Cette technique nous a permis de montrer que ParB est capable de s'étendre sur une distance d'environ vingt kilobases de part et d'autre du centromère.

Nos résultats semblent incompatibles avec le précédent modèle dans lequel ParB serait capable de polymériser sur l'ADN, à la manière d'un protéo-nucléofilament qui s'initierait au niveau du centromère. Ainsi ces travaux, nous ont permis de proposer un nouveau modèle dans lequel le complexe de partition "étendu" serait une structure concentrée, dynamique et résultant d'interactions stochastiques entre ParB et l'ADN avoisinant *ParS*.

Abstract

Segregation of genetic material over generations is an essential process ensuring that every daughter cell receives a copy of each DNA molecule. Similarly to Eukaryotes, Prokaryotes possess cytoskeletal machineries, named Par, responsible for DNA segregation. Bacterial Par systems, found on chromosomes as well as on various low copy number plasmids, are composed of three elements: a ParA protein, a ParB protein and a centromere site, *parS*. ParA ATPase is able to position plasmids in the cytoplasm during the cell cycle. Its dynamic pattern make it the motor of the partition. The centromere binding protein (CBP) ParB, binds the centromere to form a nucleoprotein assembly called the "partition complex". Using different mode of DNA binding and multiple protein-protein interactions, ParB is also capable of organizing into higher order complexes called the "extended" partition complex. This complex is the substrate for the partitioning process. Formation of the extended partition represents the first step in partition and is essential to segregation. The architecture of this complex is not known for any partition system parABS.

Here, we focus on the assembly of the F partition complex. During my PhD, I initiated several projects in parallel to characterize the different interactions involved in the organization of the partition complex and the extended partition complex of this plasmid with *in vivo* and *in vitro* approaches. In collaboration with the laboratory of Dr. Veronique Le Berre in Toulouse (LISBP -INSA), we determined *sopC* basis involved in specific SopB-*sopC* interactions. Then, we identified a new ParB determinant, outside of the helix-turn-helix DNA binding motif, responsible for specific DNA binding to the centromere. These findings reveal that ParB have an extended DNA binding domain, composed of two separate DNA binding motifs. We extended our analysis to chromosomal ParB and show that this second centromere binding motif is highly conserved in a wide range of bacteria. Using *in vivo* and *in vitro* approaches, we show that the extended partition complex architecture requires both protein-protein and protein-DNA interactions. To investigate the overall organization of the SopB-*sopC* extended partition complex, we use chromatin immunoprecipitation (ChIP) coupled with high throughput sequencing. This technique allowed us to visualize that SopB is able to extend around *sopC* over ~20 Kb.

Our results are thus inconsistent with previous models suggesting that SopB polymerize side by side in a proteo-nucleofilament emanating from the centromere. So, we propose a new model in which the extended partition complex of F plasmid assembles in a nucleoprotein complex from stochastic binding of SopB on neighboring *sopC* DNA.

Table des abréviations

| | |
|-------------|---|
| aa | Acide aminé |
| ADN | Acide DesoxyRiboNucléique |
| ADP | Adenosine Diphosphate |
| ARN | Acide RiboNucléique |
| ATP | Adenosine Triphosphate |
| ChIP | <u>Ch</u> romatin <u>I</u> mmunoprécipitation |
| Da | Dalton |
| DLS | Dynamic Light Scattering |
| dNTP | désoxyribonucléotides Tri-phosphate |
| DO | Densité Optique |
| DTT | Dithiothreitol |
| EDTA | Acide éthylène-diamine-tétraacétique |
| FISH | Fluorescent In Situ Hybridation |
| FROS | Fluorescent Repressor-Operator System |
| Fr | Fraction |
| HTH | Helix Turn Helix |
| H | heure |
| IPTG | Isopropyl β -D-1-thiogalactopyranoside |
| Kb | kilobase |
| Kd | Constante d'affinité |
| KDa | Kilodalton |
| LB | Lysogeny Broth |
| Mb | Mégabase |
| ME | Microscopie Electronique |
| min | minute |
| ml | millilitre |
| mol | molécule |
| NAPs | Nucléoïde Associated Protéins |
| <i>OriC</i> | Origine Chromosomique |
| PAGE | polyacrylamide Gel Electrophoresis |
| pb | paires de bases |
| PCR | Polymerase Chain Reaction |
| PM | Poids moléculaire |
| RHH | Ribbon-Helix-Helix |
| SDS-PAGE | Sulfate Dodécylque Sodium- PolyAcrylamide Gel Electrophoresis |
| SMC | Structure Maintenance of Chromosome |
| SPR | Surface Plasmon Resonance |
| Ter | Terminus |
| TLC | Thin Layer Chromatography |

Table des Figures

| | | |
|-----------|---|----|
| Figure 1 | Microbiologistes célèbres et leurs découvertes | 2 |
| Figure 2 | Les différents modes de répllication des réplicons bactériens circulaires | 10 |
| Figure 3 | Initiation de la répllication du chromosome d'E. coli et de différents types de plasmide | 11 |
| Figure 4 | Localisation du chromosome bactérien dans la cellule | 22 |
| Figure 5 | Ségrégation lors de la sporulation de B. subtilis | 26 |
| Figure 6 | Ségrégation et division lors de la sporulation de S. coelicolor | 27 |
| Figure 7 | Ségrégation du chromosome de C. crescentus | 28 |
| Figure 8 | Le nucléoïde bactérien | 30 |
| Figure 9 | Organisation et compaction du intra-chromosomique du chromosome bactérien | 32 |
| Figure 10 | Visualisation des boucles d'ADN formées par MatP | 34 |
| Figure 11 | FtsK, l'acteur clé de la fin de la ségrégation chez E. coli | 40 |
| Figure 12 | Diversité des organisations génétiques des opérons par plasmidiques | 44 |
| Figure 13 | Stimulation de l'activité ATPase de SopA du plasmide F d'E. coli | 52 |
| Figure 14 | Le modèle de "réaction-diffusion" | 54 |
| Figure 15 | Filaments des ParA plasmidiques et chromosomiques de type I, visualisés in vitro en ME après coloration négative | 56 |
| Figure 16 | Comportement dynamique des ParA plasmidiques et chromosomiques, in vivo | 60 |
| Figure 17 | Modèle de "pushing" du mécanisme de la ségrégation du plasmide R1 | 70 |
| Figure 18 | Structure des filaments de ParM | 72 |
| Figure 19 | Modèle de la polymérisation de ParM en "marche d'escalier" | 73 |
| Figure 20 | Appariement des plasmides R1 | 74 |
| Figure 21 | Modèle en "Tapis roulant" pour les systèmes de partition de type III -Exemple du modèle de partition du plasmide pBtoxis | 76 |
| Figure 22 | Répartition des systèmes parABS sur le chromosome principal des bactéries de l'arbre phylogénétique bactérien | 82 |
| Figure 23 | Organisation des centromères plasmidiques | 84 |
| Figure 24 | Sites centromériques sopC et parS du plasmide F et du prophage P1 | 86 |
| Figure 25 | Structures des ParB de type Ib | 88 |
| Figure 26 | Structures des protéines ParR et des complexes de partition de type II | 90 |

| | |
|--|-----|
| Figure 27 ∫ Structures des protéines TubR et des complexes de partition de type III | 94 |
| Figure 28 ∫ Organisation des domaines fonctionnels des ParB homologues codés par le plasmide F, P1 et RK2 | 96 |
| Figure 29 ∫ Structure d'un dimère de ParB (142-324) fixé ou non sur son centromère parS... .. | 98 |
| Figure 30 ∫ Structure d'un dimère de ParB (142-324) fixé ou non sur son centromère parS. | 100 |
| Figure 31 ∫ Modèle de l'étalement des protéines ParB autour du site centromérique | 108 |
| Figure 32 ∫ Appariement intramoléculaire entre deux complexes de partition portés par le plasmide pSM19035..... | 114 |
| Figure 33 ∫ Modèle de l'initiation de la polymérisation in vivo ou "DNA free zone" | 118 |
| Figure 34 ∫ Schematic representation of the F plasmid sop operon and current ParB spreading model. | 180 |
| Figure 35 ∫ Intracellular SopB proteins level and control of the sonication step for ChIP assays. | 182 |
| Figure 36 ∫ SopB binds DNA outside sopC asymmetrically. | 184 |
| Figure 37 ∫ DNA profiles from formaldehyde-treated cells show extensive sonication prior to ChIP-sequencing analysis. | 186 |
| Figure 38 ∫ Formation of secondary SopB-DNA complexes requires long DNA fragments..... | 186 |
| Figure 39 ∫ SopB binding with specific and non-specific DNA probes. | 188 |
| Figure 40 ∫ Complex B'2 is not composed of two B1 complexes..... | 188 |
| Figure 41 ∫ DNA spacing variation between two adjacent binding sites reveals that SopB dimer could bind every 16-bp without steric hindrance..... | 190 |
| Figure 42 ∫ DNase I footprinting analyses reveal no protection pattern outside SopB-binding sites. | 192 |
| Figure 43 ∫ SopB assembles into compact clusters. | 194 |
| Figure 44 ∫ Expression, activity and localization of Sop proteins in condition used for epifluorescence PALM experiments..... | 195 |
| Figure 45 ∫ The majority of SopB is recruited to the partition complexes and assembled in a small-sized nucleoprotein complex. | 196 |
| Figure 46 ∫ Model for the extended partition complex dynamic organization..... | 198 |
| Figure 47 ∫ Organisation fonctionnelle de SopB et carte des délétions construites | 218 |
| Figure 48 ∫ Profil d'élution de SopBΔ8 sur une colonne de chromatographie d'exclusion..... | 219 |

| | |
|---|-----|
| Figure 49 ∫ Activité de fixation spécifique de SopB et des variants SopB Δ 23, SopB Δ 9 et SopB Δ 8 sur sopC..... | 220 |
| Figure 50 ∫ Analyse en SPR de la fixation de SopA sur SopB ou ces variants SopB Δ 23, SopB Δ 9 et SopB Δ 8 immobilisés sur des sondes sopC | 222 |
| Figure 51 ∫ Effet des 3 variants de SopB sur la stimulation de la polymérisation de SopA | 224 |
| Figure 52 ∫ Activation de l'activité ATPase de SopA par SopB ou ces variants SopB Δ 23, SopB Δ 9 et SopB Δ 8..... | 226 |
| Figure 53 ∫ Organisation fonctionnelle de SopB et carte des délétions construites | 230 |
| Figure 54 ∫ Dernière étape de purification des protéines SopB-WT et des variants | 232 |
| Figure 55 ∫ Mesure par EMSA de la capacité des différents mutants testés à se fixer au centromère et à faire le complexe B2 | 234 |
| Figure 56 ∫ Les mutants forment des foci moins intenses et délocalisés..... | 236 |
| Figure 57 ∫ Etalement de SopB, SopB-R ₁₀₈ A et SopB-11 autour du centromère sopC | 238 |
| Figure 58 ∫ Répartition de l'étalement de SopB selon les réplicons : chromosomique ou plasmidiques | 242 |

Liste des tableaux

| | |
|---|-----|
| Tableau 1 ∫ Acteurs impliqués dans la ségrégation des chromosomes bactériens | 35 |
| Tableau 2 ∫ Diversité des nomenclatures et principales caractéristiques des systèmes de ségrégation | 44 |
| Tableau 3 ∫ Variations autour des conditions de polymérisation entre les différents systèmes de type I | 58 |
| Tableau 4 ∫ Caractéristiques et variations des séquences centromériques..... | 80 |
| Tableau 5 ∫ Analyse conformationnelle de la partie N-terminale de SopB et de ces variants en fonction de leur comportement en chromatographie d'exclusion. | 219 |
| Tableau 6 ∫ Plasmides utilisés dans cette étude | 275 |
| Tableau 7 ∫ Oligonucléotides utilisés pour l'expérience de gel EMSA et de Footprint..... | 276 |
| Tableau 8 ∫ Oligonucléotides utilisés pour l'expérience de qChIP..... | 277 |
| Tableau 9 ∫ Oligonucléotides utilisés pour la construction des plasmides de cette étude | 278 |
| Tableau 10 ∫ Souches utilisées dans cette étude..... | 279 |

Table des matières

PARTIE I - INTRODUCTION 23

| | |
|--|----|
| I. COMPOSITION ET MAINTIEN DU GENOME BACTERIEN | 3 |
| A. La composition du génome bactérien | 3 |
| 1. Le chromosome..... | 3 |
| 2. Les plasmides | 5 |
| 3. Les bactéries multichromosomiques : différences entre plasmides et chromosomes..... | 7 |
| 1. Les acteurs du maintien des réplicons | 13 |
| B. La ségrégation des plasmides et ses systèmes de contrôle | 15 |
| 1. La ségrégation des plasmides..... | 15 |
| a) La ségrégation par diffusion passive | 17 |
| b) La ségrégation active | 17 |
| 2. Les systèmes de secours | 21 |
| C. La ségrégation des chromosomes bactériens | 21 |
| 1. Ségrégation active des origines..... | 23 |
| 2. Ségrégation du corps du chromosome | 31 |
| a) Organisation et compaction du chromosome - (les complexes SMC) | 31 |
| b) L'expulsion par le réplisome et l'implication de la transcription..... | 37 |
| c) L'entropie | 39 |
| 3. La ségrégation du terminus : FtsK/SpoIII E..... | 41 |
| II. MECANISMES MOLECULAIRES DES SYSTEMES DE PARTITION | 45 |
| A. Les systèmes de partition de type I | 47 |
| 1. L'activité ATPase ParA et autorégulation de l'opéron | 51 |
| 2. Le comportement dynamique de ParA | 57 |
| a) Les polymères de ParA..... | 57 |
| (1) Rôle de l'ATP dans la polymérisation de ParA..... | 59 |
| (2) Rôle de ParB et de l'ADN dans la polymérisation de ParA | 59 |
| b) L'oscillation de ParA..... | 61 |
| B. Les systèmes de partition de type II..... | 69 |
| C. Les systèmes de partition de type III..... | 77 |
| III. LE COMPLEXE DE PARTITION | 79 |
| A. Le centromère nature et diversité | 79 |
| 1. Les centromères chromosomiques | 81 |

| | |
|---|------------|
| 2. Les centromères plasmidiques..... | 85 |
| B. Les protéines adaptatrices ou de liaison au centromère : Nature et diversité..... | 89 |
| C. Structure et formation des complexes de partition..... | 91 |
| 1. Le complexe de partition de type II..... | 91 |
| 2. Le complexe de partition de type III..... | 95 |
| 3. Les complexes de partition de type I | 97 |
| D. Le complexe de partition "étendu" pour le type I | 105 |
| 1. Etude topologique du complexe de partition | 107 |
| 2. L'étalement de ParB sur l'ADN | 109 |
| 3. Rôle du complexe de partition "étendu" | 115 |
| a) L'appariement des plasmides | 115 |
| b) L'interaction entre le complexe de partition et la protéine ParA..... | 119 |
| PARTIE II - RESULTATS | 125 |
| CHAPITRE I..... | 131 |
| I. ETAT DE L'ART | 131 |
| II. PRESENTATION DE L'ARTICLE : | 132 |
| CHAPITRE II..... | 155 |
| I. ETAT DE L'ART | 155 |
| II. PRÉSENTATION DE L'ARTICLE: "INSIGHT INTO CENTROMERE-BINDING PROPERTIES OF PARB" | 156 |
| CHAPITRE III..... | 177 |
| I. ETAT DE L'ART | 177 |
| II. PRÉSENTATION DE L'ARTICLE : "THE BACTERIAL SEGROSOME ASSEMBLES WITHOUT REQUIRING PARB FILAMENT FORMATION" | 179 |
| A. Abstract..... | 179 |
| B. Introduction | 179 |
| C. Results | 183 |
| 1. SopB propagates over large region on both F and mini-F plasmids | 183 |
| 2. SopB do not propagate linearly along the DNA outside the centromere | 185 |
| 3. RepE/iteron nucleoprotein complexes limit SopB spreading without acting as a barrier | 187 |
| 4. Formation of the extended partition complex requires both protein-protein and protein-DNA interactions | 187 |
| 5. SopB dimers assemble every 16-bp along the DNA..... | 189 |
| 6. Extended partition complexes do not assemble in stable nucleoprotein complex | 191 |

| | |
|--|-------------------|
| 7. All SopB molecules are concentrated in foci at centromere sites | 193 |
| 8. Discussion..... | 199 |
| D. Material and methods | 202 |
| 1. Bacterial strains, plasmids and growth conditions | 202 |
| 2. Chromatin immuno-precipitation | 203 |
| 3. q-ChIP analysis..... | 204 |
| 4. ChIP-seq..... | 204 |
| 5. Electromobility shift assay | 204 |
| 6. Footprint assays | 204 |
| 7. PhotoActivated Localization Microscopy and epifluorescence microscopy | 205 |
| E. Bibliography..... | 206 |
| F. Supplementary data | 210 |
| CHAPITRE IV | 215 |
| I. ETAT DE L'ART | 215 |
| II. CARACTERISATION DU DOMAINE D'INTERACTION ENTRE SOPB ET L'ATPASE SOPA | 217 |
| A. Approche de mutagenèse par délétion: Construction des 3 mutants, SopB Δ 23, SopB Δ 9 et SopB Δ 8 | 217 |
| B. Purification et implication de la partie N-terminale de SopB dans le comportement non globulaire de la protéine | 217 |
| C. Interaction de SopB et de ces variants avec le centromère, <i>sopC</i> | 221 |
| D. Mesure d'interaction de SopA avec SopB ou les protéines SopB tronquées | 223 |
| E. Effet des variants de SopB sur la stimulation de la polymérisation de SopA..... | 225 |
| F. Effet des variants de SopB sur la stimulation de l'activité ATPase de SopA | 227 |
| G. Conclusion-Discussion | 228 |
| III. CARACTERISATION DU DOMAINE D'OLIGOMERISATION DE SOPB..... | 231 |
| A. Approche de mutagenèse par délétion et mutation ponctuelle | 231 |
| B. Profil de fixation sur un ADN contenant le site centromérique <i>sopC</i> | 233 |
| C. Localisation des différents variants en microscopie à fluorescence..... | 237 |
| D. Analyse en qChIP de la capacité des différents variants de SopB à réaliser de l'étalement | 239 |
| E. Discussion..... | 240 |
| IV. LA PROPAGATION DES PROTEINES SOPB SUR LE CHROMOSOME D' <i>E. COLI</i> | 243 |
| V. PERSPECTIVES..... | 244 |
| <u>PARTIE III - DISCUSSION GENERALE</u> | <u>249</u> |

| | |
|---|------------|
| PARTIE IV - MATERIELS ET METHODES | 259 |
| I. PLASMIDES, SOUCHES, MUTAGENESE ET ADN | 261 |
| A. Mutagénèse par délétion..... | 261 |
| B. Mutagénèse dirigée réalisée par la technique de "quick change" | 261 |
| C. Mutagénèse avec le kit "In fusion" | 261 |
| II. MILIEUX ET ANTIBIOTIQUES UTILISES..... | 262 |
| III. TRANSFORMATION | 262 |
| IV. EXPRESSION ET PURIFICATION DE SOPA ET SOPB | 263 |
| A. SopA | 263 |
| B. Purification de SopB étiquetée intéine | 264 |
| V. TEST DE RETARD DE MIGRATION SUR GEL..... | 265 |
| A. Préparation des sondes..... | 265 |
| A. Réactions d'interaction <i>in vitro</i> | 266 |
| VI. WESTERN-BLOT..... | 266 |
| VII. FOOTPRINT <i>IN VITRO</i> | 267 |
| VIII. TEST D'ACTIVITE ATPASE..... | 268 |
| IX. TEST DE STABILITE | 268 |
| X. MICROSCOPIE A FLUORESCENCE (CHAPITRE IV)..... | 269 |
| XI. RESONNANCE PLASMONIQUE DE SURFACE | 269 |
| XII. EXPERIENCE DE DIFFUSION DE LUMIERE DYNAMIQUE (DLS) | 270 |
| XIII. IMMUNOPRECIPITATION DE CHROMATINE (CHIP)..... | 270 |
| A. Purification des anticorps anti-SopB..... | 270 |
| B. Culture et traitement au formaldéhyde..... | 271 |
| 1. Sonication..... | 271 |
| 2. Immunoprécipitation et purification de l'ADN..... | 272 |
| XIV. PCR QUANTITATIVE..... | 273 |
| A. Choix des différentes amorces..... | 273 |
| B. Réaction de PCR quantitative..... | 273 |
| C. Analyse des réactions de PCR quantitative | 273 |
| XV. SEQUENÇAGE HAUT DEBIT | 274 |
| PARTIE V - BIBLIOGRAPHIE..... | 281 |

Introduction

a

b

c

d



Figure 1 \int Microbiologistes célèbres et leurs découvertes

- a) Theodor Escherich (1857-1911) était un pédiatre et bactériologiste allemand-autrichien. Il découvrit en 1885, la célèbre bactérie qui porte son nom, *Escherichia coli*.
- b) Image de la bactérie *E. coli* en Microscopie électronique (ME). *E. coli* est une bactérie intestinale commensale à Gram négatif. Elle compose environ 80 % de notre flore intestinale. Certaines souches d'*E. coli* peuvent être pathogènes entraînant alors des gastro-entérites, infections urinaires, méningites ou septicémies. Mais, elle est aussi une bactérie modèle de laboratoire pour l'étude des procaryotes.
- c) Joshua Lederberg (1925-2008) est un généticien et microbiologiste américain, pionnier de la biologie moléculaire. Il a reçu le prix Nobel de médecine en 1958 avec Edward Tatum et George Wells Beadle pour avoir découvert l'existence d'échange de matériel génétique entre bactéries, appelé la conjugaison bactérienne (**Encadré 1**). Il découvre ensuite, par l'étude des bactéries *Salmonella*, les plasmides. Les plasmides sont des morceaux d'ADN présents dans les cellules bactériennes qui se répliquent de manière autonome, séparément de l'ADN chromosomique.
- d) Image d'un plasmide circulaire en ME

I. Composition et maintien du génome bactérien

A. La composition du génome bactérien

Le génome d'une cellule est essentiel à son identité : Il contient l'ensemble des déterminants génétiques définissant les capacités et les limites intrinsèques de chaque espèce donnée. Au cours du cycle cellulaire, le génome doit être dupliqué par réplication puis correctement ségrégué dans les futures cellules filles afin d'assurer le maintien de l'information génétique au cours des générations. Chez les bactéries, le génome est constitué de deux grands types de molécules d'ADN : les chromosomes et les éléments extra-chromosomiques appelés plasmides. Ces deux types de molécules sont capables de se répliquer de manière autonome dans la cellule et ont été appelées par Jacob et Brenner, réplicons (JACOB and BRENNER, 1963).

1. Le chromosome

Les chromosomes sont généralement des molécules circulaires de grande taille (de l'ordre du Mb) comme l'avait défini (JACOB and WOLLMAN, 1956)), et porteurs de gènes essentiels au développement de la bactérie (en particulier les gènes ribosomiques). Cependant aucun de ces arguments ne peut rigoureusement servir à le définir. Certaines bactéries comme *Streptomyces* ou *Borrelia* ont un chromosome linéaire alors que la grande majorité des autres, possède un chromosome circulaire. La bactérie la plus étudiée, avec son grand chromosome circulaire d'environ 4.6Mb est *Escherichia coli* (**Figure 1**). La conversion de son chromosome circulaire en chromosome linéaire ne semble pas avoir de préjudice pour la biologie de la cellule (Cui *et al.*, 2007; Chaconas and Kobryn, 2010). Ainsi, dans un même organisme, les processus d'organisation du chromosome, de réplication et de ségrégation semblent être indépendants du fait que le chromosome soit circulaire ou linéaire. D'autre part, les chromosomes possèdent des tailles variables allant de 160 kilobases (Kb) pour le plus petit chromosome identifié (*Carsonella ruddii*), à 13 mégabases (Mb) pour le chromosome de *Sorangium cellulosum*. La taille du génome peut refléter le mode de vie de la bactérie : il est souvent de plus grande taille pour les bactéries ayant un métabolisme complexe et/ou différentes étapes de développement comme la sporulation ou la formation de mycélium, tandis qu'il sera de petite taille pour les bactéries vivant en symbiose ou comme parasites obligatoires.

Les chromosomes sont autonomes pour leur réplication. La réplication des chromosomes circulaires est bidirectionnelle et s'initie spécifiquement à une seule origine de réplication chromosomique appelée *OriC*. Les régions proches d'*OriC* sont répliquées en premier alors que les régions proches du terminus le seront ultérieurement. Dans les cellules en cours de réplication, le nombre de copies des loci proches d'*OriC* est par conséquent supérieur au nombre de loci proche du terminus. Lors de la croissance rapide, en milieu riche, la majorité des bactéries comme *E. coli* sont capables d'initier des cycles réplcatifs chevauchants qui pourront s'achever au cours de la génération suivante (Cooper and Helmstetter, 1968). Ainsi, le nombre de copies des gènes proches des origines peut varier d'un facteur 2 à 16 selon le type de milieu (pauvre ou riche) et les conditions de croissance, ce qui est induit naturellement un dosage des gènes en fonction de leur localisation sur le chromosome (Couturier and Rocha, 2006).

Le mécanisme de réplication définit deux bras de réplication appelés réplicores et consiste en la synthèse continue d'un des deux brins (brin meneur ou Leading) du chromosome et discontinue de l'autre (brin retardé ou Lagging). La composition en base du chromosome est biaisée selon les deux réplicores (Blattner, 1997). Ce biais permet de déterminer l'origine et la terminaison de la réplication puisque c'est au niveau de ces deux régions que le biais GC s'inverse (Lobry, 1996; Rocha *et al.*, 2006). Ce biais concerne aussi plusieurs motifs d'oligonucléotides qui ont des fonctions spécifiques selon les deux réplicores. Par exemple, l'orientation biaisée des séquences KOPS, permet à l'ADN translocase FtsK de s'orienter en direction des sites *ter* situés dans la région terminus. Au niveau de ces sites, FtsK pourra recruter d'autres protéines impliquées dans le processus de ségrégation des termini et de résolution des dimères de chromosome (Bigot *et al.*, 2007; Stouf *et al.*, 2013).

2. Les plasmides

En microbiologie, un plasmide désigne un élément génétique extrachromosomique, physiquement distinct de l'ADN chromosomique qui est capable de réplication autonome. Ce terme *plasmide* fut introduit par le biologiste moléculaire américain J. Lederberg (LEDERBERG, 1952) (**Figure 1**). Typiquement, les plasmides sont des molécules de petites tailles et présents en multi-copie (<10 Kb ; >15 copies/ chromosome), ou de grandes tailles (> 30 Kb) et présents en faible nombre de copies. La gamme de taille des plasmides peut donc considérablement

Encadré 1- Le maintien des plasmides dans quel but ?

Les plasmides peuvent être porteurs de gènes de résistance aux antibiotiques ou aux métaux lourds, de gènes permettant l'utilisation de certains nutriments utiles pour la croissance sur sols pollués. Ils peuvent également être porteurs de facteurs de virulence comme le complexe pili-adhésine chez *Salmonella*, permettant l'adhésion aux entérocytes eucaryotes ou encore de gènes de production de différentes toxines appelés colicines pouvant être létales pour les autres bactéries environnantes.

Les plasmides participent également aux transferts horizontaux de gènes entre les populations bactériennes, et donc à la dissémination des gènes conférant des avantages sélectifs (par exemple des résistances aux antibiotiques ou des facteurs de virulence). La mobilité des plasmides (par conjugaison) au sein des populations bactériennes accroît le spectre d'hôte des gènes impliqués dans la virulence. Ces gènes offrent en contrepartie un avantage sélectif pour le plasmide et les bactéries hôtes. On conçoit donc la nature quasi-ubiquitaire et persistante des plasmides chez les bactéries pathogènes.

Encadré 2- Les transferts horizontaux

Des analyses comparatives des génomes procaryotes montrent que l'acquisition de matériel génétique par transfert horizontal de gènes a été une force motrice majeure dans l'évolution du règne bactérien (Koonin *et al.*, 2001). Les transferts horizontaux s'effectuent via trois mécanismes connus à l'heure actuelle : la transduction, la transformation et la conjugaison.

- a) **La transduction** : Le transfert d'ADN chromosomique ou d'ADN plasmidique est réalisé par des bactériophages (pour revue Dreiseikelmann, 1994). L'exemple le plus connu est le pro-phage P1 qui existe sous forme de plasmide chez *E. coli*, le plasmide P1, pendant sa phase de lysogénie.
- b) **La transformation** : C'est un mécanisme par lequel de l'ADN exogène est internalisé dans le cytoplasme et intégré dans le génome par recombinaison homologe si sa séquence présente suffisamment d'homologie avec le chromosome.
- c) **La conjugaison** est un processus durant lequel deux bactéries vont s'échanger de l'information génétique. Pour cela, une bactérie dite donneuse met en place un pilus de conjugaison avec une bactérie dite réceptrice, permettant le passage d'ADN simple brin entre les deux bactéries. Le plasmide F d'*E.coli*, aussi appelé facteur F pour facteur de fertilité, est le premier exemple de conjugaison décrit. Il contient toute l'information nécessaire à son propre transfert par conjugaison et à son maintien. C'est un "épisode" capable de s'intégrer au chromosome, ou s'exciser sous une forme plasmidique circulaire d'une taille d'environ 100 kpb (Scaife 1966).

Les plasmides utilisent donc deux modes de transfert horizontaux: La transduction et la conjugaison.

varier, allant de quelques Kb à plusieurs Mb, comme pour le mégaplasme de la bactérie du sol, *Ralstonia solanacearum* GMI100 (2.09 Mpb) (Salanoubat *et al.*, 2002) ou de *Cupriavidus metallidurans* CH34 (2.58 Mb). Des analyses statistiques montrent qu'il existe une corrélation entre la taille des plasmides et la pression écologique que subissent les bactéries qui en sont porteuses (Slater *et al.*, 2008). En effet, bien qu'étant pour la plupart non-essentiels à la survie de la bactérie, ils sont souvent porteurs de déterminants génétiques conférant un avantage sélectif à la bactérie hôte (**voir encadré 1**). Pour cela, les plasmides sont maintenus dans une population soit par transfert horizontal soit par transfert vertical. Les différentes formes de transferts horizontaux sont décrites dans l'**encadré 2** et le transfert vertical sera décrit plus loin.

3. Les bactéries multichromosomiques : différences entre plasmides et chromosomes

La grande majorité des bactéries ne portent qu'un seul chromosome. Cependant, on découvre de plus en plus de bactéries avec deux ou trois grands réplicons comme *Agrobacterium tumefaciens*, *Vibrio cholerae* ou encore *Burkholderia cenocepacia*. Le principal critère utilisé pour différencier un chromosome d'un plasmide est la présence de gènes essentiels. Chez ces bactéries, le chromosome le plus grand contenant la majorité des gènes essentiels (gènes de ménages et gènes codants les ARN ribosomiques) et possédant une origine de réplication similaire aux bactéries mono-chromosomiques est appelé, chromosome principal. A l'inverse, les autres chromosomes ne portant qu'une petite partie des gènes essentiels, sont appelés chromosomes secondaires. L'initiation de la réplication du chromosome principal et des chromosomes secondaires est synchronisée avec le cycle cellulaire et arrive à des moments précis. Au contraire, l'initiation de la réplication des plasmides peut avoir lieu à tout moment du cycle cellulaire. Les chromosomes secondaires possèdent des caractéristiques hybrides entre les chromosomes principaux et les plasmides. Leur région origine est relativement variable à l'image de celle observée chez les plasmides et peut être identifiée par leur structure (Egan and Waldor, 2003), ou par leur homologie de séquence avec les systèmes plasmidiques connus (Dubarry *et al.*, 2006). De même, les gènes présents sur les chromosomes principaux sont très conservés alors que les gènes présents sur les chromosomes secondaires ont une forte variabilité évolutive comme les plasmides (Cooper *et al.*, 2010). La composition en bases d'un réplicon, calculé par le pourcentage en Guanine et Cytosine (%GC) est un bon indicateur de cette plasticité évolutive. Le pourcentage GC est

homogène sur un réplicon donné et varie d'une espèce à l'autre. Ainsi un fragment d'ADN (îlot génomique) ou un plasmide, acquis récemment par transfert horizontal, présentent parfois un % GC différent des chromosomes primaires et secondaires pour lesquels le % GC reste très proche. On retrouve davantage de traces d'échanges horizontaux sur les chromosomes secondaires que sur les chromosomes primaires. Au vu de cette plasticité évolutive, les chromosomes secondaires pourraient servir "de réservoir de diversité ou de banc de test de l'évolution".

Le maintien de l'état multi-chromosomique pose la question de l'avantage sélectif à l'origine de ce maintien. Plusieurs hypothèses sont avancées comme la réduction du temps de génération du génome ou le phénomène de dosage de gènes. En effet, pour une taille équivalente, un génome mono-chromosomique, répliqué à partir d'une seule origine de réplication, sera plus long à dupliquer qu'un génome multi-chromosomique dont les différents chromosomes seront répliqués en parallèles. Il peut aussi exister un phénomène de dosage de gènes entre le chromosome principal et les chromosomes secondaires, similaire à celui qui existe entre les régions proches de l'origine de réplication (*OriC*) et celles proches du terminus de réplication (*ter*) sur le chromosome principal. Par exemple chez *V. cholerae*, le décalage des initiations de réplication entre les deux chromosomes entraîne une diminution d'un facteur quatre de l'expression des gènes présents sur le chr2 par rapport à ceux proches de l'*OriC* du chr1. Les auteurs suggèrent que ceci pourrait permettre de conserver un ensemble important de gènes faiblement exprimés (de fonctions inconnues) (Srivastava and Chattoraj, 2007). Ainsi la présence de plusieurs chromosomes pourraient augmenter l'expression d'un plus grand nombre de gènes nécessaires à la croissance rapide, situés proches de l'*OriC* du Chr1 (comme les gènes de ménage) et permettre de conserver un ensemble de gènes faiblement exprimés (parmi lesquels éventuellement des gènes délétères à la croissance cellulaire à plus haute expression) en les mettant proches de la région terminus du Chr2.

Bien que les avantages des bactéries multi-chromosomiques soient très difficiles à mettre en évidence en conditions de laboratoire, il est néanmoins possible qu'un faible avantage observé au laboratoire puisse suffire à améliorer la compétitivité de ces bactéries dans leur milieu de vie.

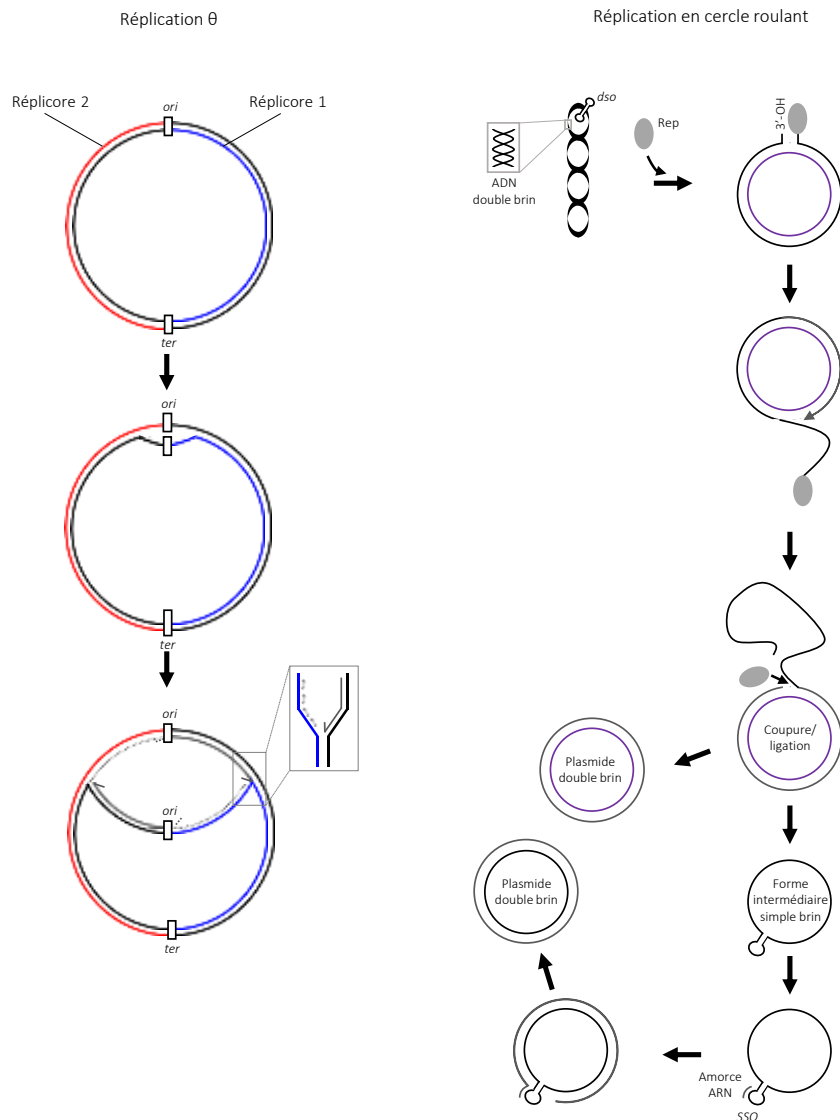


Figure 2 | Les différents modes de répllication des réplicons bactériens circulaires

La direction de répllication est indiquée par les flèches. L'ADN polymérase ne synthétise l'ADN que dans le sens 5'>3', et les deux brins d'ADN sont appariés en anti-sens. Les brins nouvellement synthétisés sont en gris (le brin leading en trait continu, le brin lagging en pointillé).

- Réplication de type θ :** La répllication est initiée au site unique *ori*. Les sites *ori* et *ter*, sont signalés et représentés par des rectangles noirs. Elle est bidirectionnelle ce qui définit deux bras de répllication appelés réplicores 1 et 2. La synthèse du brin lagging est discontinue et se fait sous forme de petits fragments, appelés fragments d'Okasaki (petites flèches). La synthèse du brin leading est quant à elle continue. Ceci permet la synthèse simultanée des deux brins. La répllication s'achève dans la région *ter*.
- Réplication en cercle roulant :** La répllication en cercle roulant est unidirectionnelle. La protéine Rep, codée par le plasmide produit une coupure dans le site *dso* du plasmide surenroulé. Rep reste accroché à l'extrémité 5'. L'extrémité 3'-OH est reconnue par la polymérase III, qui réalise la répllication en se servant du second brin intact comme matrice. Le brin parental qui est lié à Rep est séparé de son brin complémentaire et déplacé grâce à l'action d'une hélicase. La première étape de la répllication s'achève au moment où l'ADN polymérase atteint à nouveau le site *dso*. Cette étape produit une copie double brin du plasmide (composé d'un brin parental et d'un brin néosynthétisé) et une forme intermédiaire circulaire simple brin. Cette dernière est convertie en ADN double brin à la suite d'un processus répllicatif qui s'initie dans le site *sso* à partir d'une amorce synthétisée par l'ARN polymérase. Pour finir, tous les plasmides ainsi synthétisés sont surenroulés négativement par la gyrase.
- Le déplacement de brin n'est pas illustré dans cette étude (Scherzinger *et al.*, 1991; Sakai and Komano, 1996).

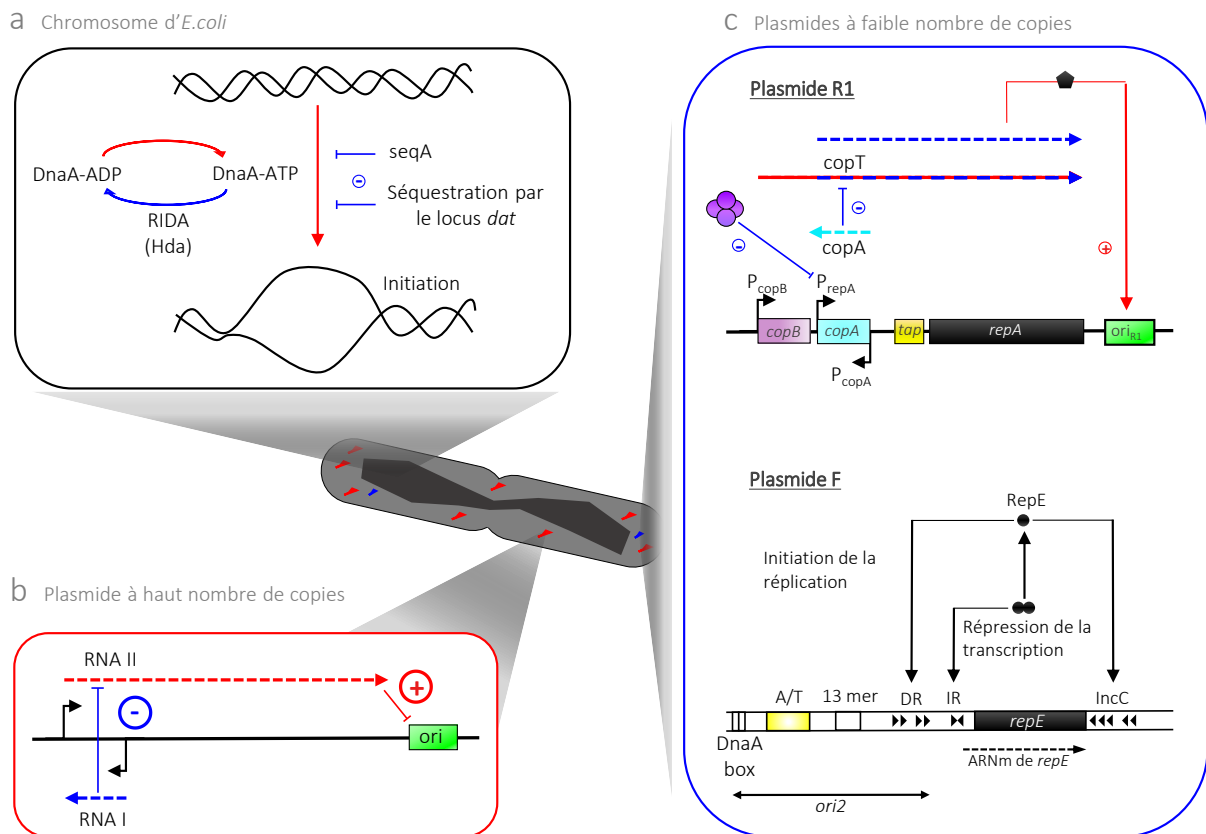


Figure 3 | Initiation de la réplication du chromosome d'*E. coli* et de différents types de plasmide

- a) **Chez *E. coli*, l'acteur majeur de la régulation est la protéine DnaA.** DnaA est active sous sa forme ATP et la fixation de DnaA-ATP sur la région riche en A/T de l'origine facilite l'ouverture de l'ADN double brin et l'initiation de la réplication. La machinerie de réplication composée des hélicases DnaB, DnaC et de la polymérase III, est alors chargée sur l'origine de réplication ouverte afin de commencer la réplication. Si DnaA-ATP se fixe sur les *OriC* à peine dupliquées, il y aura réinitiation immédiate. 3 systèmes empêchent cette réinitiation : le système RIDA (Regulatory Inactivation of DnaA), la séquestration des origines par SeqA et la titration de DnaA par le locus *data*. Dans le premier système, l'hydrolyse de DnaA-ATP est stimulée par le clamp et une protéine annexe Hda. La réplication semi-conservative est à l'origine d'une séquence double brin hémiméthylé contenant des sites GATC, substrats de SeqA. Dans le deuxième système, les origines à peine dupliquées sont recouvertes de SeqA ce qui empêche la fixation de DnaA-ATP sur la région riche en A/T. Et pour finir, le locus *data* permet de titrer DnaA en fixant 5 fois plus de DnaA que ceux présents à *OriC*, complétant ainsi la limitation de la ré-initiation.
- b) **Les plasmides à haut nombre de copies.** La réplication des plasmides de type ColE1, requiert la transcription d'un ARN (ARN II) dont l'extrémité 3' est impliquée dans un hybride ADN-ARN II reconnu par la RNaseH d'*E. coli*. Le clivage du duplex hybride par la RNase H libère une extrémité 3' hydroxyle qui sert d'amorce pour l'initiation de la synthèse d'ADN polymérase I. Le contrôle du nombre de copies s'opère par l'appariement d'un petit ARN anti-sens (ARN I), complémentaire de la partie 5' d'ARN II, ce qui va déstabiliser l'hybride ADN-ARN II reconnu par la RNase H, et par conséquent inhibe la réplication (pour revue (Nordström and Wagner, 1994)).
- c) **Les plasmides à faible nombre de copies.** De nombreux plasmides utilisent des protéines initiatrices spécifiques des plasmides, les protéines Rep. La liaison coopérative d'un grand nombre de protéines Rep au niveau de l'origine de réplication, permet l'ouverture de l'ADN double brin et le recrutement de DnaB. Deux stratégies de régulation de l'initiation sont représentées pour ces plasmides. **Dans le cas du plasmide R1**, le contrôle se fait grâce à un ARN anti-sens (pour revue (Wagner and Simons, 1994; Nordström and Wagner, 1994)). Ce contrôle est exercé sur la synthèse des protéines Rep_{R1} (RepA) au niveau transcriptionnel et post-transcriptionnel par les produits des gènes *copA* et *copB* respectivement. CopB réprime la transcription de RepA en agissant sur son promoteur. L'ARN anti-sens *copA* inhibe la traduction de l'ARNm de *repA* en se fixant à la région "leader" (CopT) de l'ARNm de *repA* (Blomberg *et al.*, 1992). **Dans le cas du plasmide F ou P1**, la protéine RepE est un homodimère inactif pour l'initiation de la réplication, mais actif pour la répression de la transcription du gène *repE*. L'initiation de la réplication nécessite la conversion des dimères en

monomères. Ces derniers se fixent sur les itérons d'*ori2* afin d'initier la réplication. Le contrôle de la réplication se fait grâce à la présence d'itérons (*IncC*) située en aval du gène *repE* qui titrent les protéines, réduisant ainsi la fréquence d'initiation. Le modèle de "Handcuffing" est une variante de ce modèle de titration (McEachern *et al.*, 1989; Morrison and Chattoraj, 2004)). Selon ce modèle, les protéines Rep fixées sur les itérons d'une origine vont interagir avec des complexes similaires d'autres origines ce qui est à l'origine de l'appariement des plasmides. Cet appariement cause un encombrement stérique au niveau des deux origines, ce qui va défavoriser l'initiation de la réplication.

1. Les acteurs du maintien des réplicons

Le maintien des réplicons au cours du cycle cellulaire dépend de deux grands processus : La réplication et la ségrégation.

Chez les chromosomes, la réplication est bidirectionnelle alors que chez les plasmides elle peut être uni ou bidirectionnelle. Trois modes de réplication ont été décrits pour les réplicons bactériens : la réplication de type θ (Thêta), la réplication par déplacement de brin et la réplication en cercle roulant qui a également lieu lors de la conjugaison de certains plasmides (Del Solar *et al.*, 1998; Chattoraj, 2000; Messer, 2002; Khan, 2005) (**Figure 2**). Le mode Thêta est utilisé par les chromosomes et les plasmides alors que les deux autres modes ne sont utilisés que par les plasmides. Le maintien d'un réplicon passe tout d'abord par le contrôle de son nombre de copies. Tous les réplicons sont capables de "mesurer" naturellement leur propre nombre de copies et ajustent en conséquence leur fréquence de réplication. Pour s'assurer que chaque future cellule fille hérite d'un réplicon, le niveau de l'initiation de la réplication doit être finement régulé (**Figure 3**). Cette régulation fait intervenir des protéines initiatrices (tels que DnaA pour les chromosomes principaux, ou Rep pour certains plasmides) ou des petits ARN régulateurs (plasmide ColE1, R1). Ces mécanismes de régulation sont détaillés dans la revue de Reyes-Lamothé et collaborateurs (Reyes-Lamothé *et al.*, 2012).

La régulation de l'initiation de la réplication des chromosomes repose sur une réduction de la disponibilité et/ou de l'activité de la protéine initiatrice DnaA et de la région origine (Pour plus de détails, voir **Figure 3**). Pour les plasmides, l'initiation de la réplication est indépendante de celle du chromosome et du cycle cellulaire. Le nombre de copies d'un plasmide peut néanmoins varier suivant la souche bactérienne hôte, les conditions de croissance et naturellement, selon le stade du cycle cellulaire. Le plasmide utilise la machinerie de réplication de l'hôte, par conséquent, son nombre de copies doit être finement régulé pour ne pas interférer avec le métabolisme de l'hôte (nombre de copies élevé) ou être perdu (faible nombre de copies). Les plasmides ont mis au point des boucles de régulations négatives, pour diminuer la disponibilité des acteurs de l'initiation. Deux stratégies ont été décrites : la première permet de réguler le niveau intracellulaire de la protéine Rep par titration avec des séquences de fixation spécifiques, appelées itérons. La deuxième inhibe la transcription d'un ARN essentiel à l'initiation de la réplication par fixation d'un ARN antisens (**Figure 3**). Les plasmides F et P1

utilisent le premier mode de fonctionnement (Chattoraj, 2000) et les plasmides ColE1 et R1 utilisent le second (Del Solar and Espinosa, 2000).

Le maintien des différents réplicons dans une même cellule, pose un problème d'incompatibilité: entre chromosomes primaires et secondaires, entre plasmides et chromosomes ainsi qu'au niveau des plasmides entre eux. Cette incompatibilité réside à deux niveaux : la réplication et la ségrégation. En effet, deux réplicons possédant le même système de réplication ou de ségrégation, tendent à s'exclure l'un l'autre et ne peuvent pas se maintenir dans une même cellule. Dans le cas de la réplication des plasmides, les systèmes de contrôle de la réplication ne pourront pas différencier les deux plasmides, et vont se comporter comme si il n'y avait qu'un seul et même plasmide, présent en deux fois plus de copies, conduisant ainsi à l'inhibition de la réplication. Le faible nombre de copies de chacun des plasmides en résultant, conduira rapidement à leur perte durant les générations suivantes. De même, pour la ségrégation, si deux réplicons portent des systèmes de ségrégation identiques ou proches, alors leurs systèmes ne pourront pas les distinguer. Fréquemment, il s'ensuit une séparation des deux réplicons plutôt qu'à celle des copies de chacun des réplicons, conduisant rapidement à l'apparition de deux populations, chacune porteuse de l'un ou l'autre des plasmides. Ainsi pour pouvoir subsister dans une même cellule, deux réplicons devront posséder des systèmes de réplication et de ségrégation différents. La ségrégation des plasmides et des chromosomes sera détaillée tour à tour dans les sections suivantes.

B. La ségrégation des plasmides et ses systèmes de contrôle

1. La ségrégation des plasmides

La problématique du maintien d'un plasmide dans une population est très différente selon sa taille et son nombre de copies dans la cellule bactérienne. On distingue :

- Les plasmides à bas nombre de copies, comme le plasmide F ou P1 d'*E. coli*, présent à 1-5 copies par cellule selon les conditions de croissance
- Les plasmides à haut nombre de copies, comme le plasmide ColE1, présent entre 20 et 30 copies par cellule

a) La ségrégation par diffusion passive

La diffusion passive des plasmides dans le cytoplasme engendre une répartition aléatoire lors de la division cellulaire. Les plasmides à haut nombre de copies comme les plasmides ColE1 (quelques dizaines de copies/chromosome), peuvent se maintenir par diffusion passive au cours des générations bactériennes. La probabilité d'apparition d'une cellule sans plasmide, par division cellulaire, est prédite par la formule suivante : $P_0=2^{1-n}$, où n est le nombre de copies du plasmide avant la division cellulaire. La perte d'un plasmide présent en 10 exemplaires est prédite toutes les 512 générations, ce qui est peu fréquent. Par contre, un plasmide présent à deux copies par cellule sera perdu toutes les deux générations.

Bien qu'étant un contributeur clé des mouvements des macromolécules, la diffusion passive ne peut maintenir à elle seule, les plasmides à faible nombre de copies, comme le plasmide F. Ainsi, les grands plasmides, présents en faible nombre de copies, ont dû acquérir des systèmes de ségrégation actifs pour se maintenir dans les cellules bactériennes.

b) La ségrégation active

Les plasmides à bas nombre de copies possèdent donc des systèmes de partition actifs, appelées *parABS* ou locus *par*. Ce locus est composé de trois éléments : un segment d'ADN appelé "séquence centromérique" ou *parS* et de deux gènes *parA* et *parB* codant respectivement une protéine NTPase, impliquée dans le positionnement du plasmide lors de la partition et une protéine qui se fixe spécifiquement sur cette séquence. Les premières analyses ont été effectuées sur le plasmide F d'*E. coli* et ont permis de définir deux régions importantes pour le maintien du plasmide : une région essentielle à la réplication (Timmis *et al.*, 1975) et une région essentielle à la stabilité du plasmide (Bergquist *et al.*, 1981). Cette dernière est capable de stabiliser des plasmides instables porteurs de leur origine de réplication sans augmenter leur nombre de copies (Ogura *et al.*, 1980; Ogura and Hiraga, 1983). L'analyse par délétions successives et par des expériences de complémentation de cette région essentielle à la stabilité du plasmide F, a permis de mettre en évidence le premier locus de partition, appelé spécifiquement *sopABC* (pour stability of plasmid). Ce locus est composé deux gènes nommés *sopA* et *sopB*, et une région nommée *sopC* qui est la séquence centromérique. Le même type de locus, appelé *parABS*, a été retrouvé sur le plasmide P1 d'*E. coli* (Austin and Abeles, 1983; Abeles *et al.*, 1985). Ces systèmes ont donc initialement été mis en évidence sur les plasmides

F et P1 d'*E. coli* (Ogura and Hiraga, 1983; Austin and Abeles, 1983), puis leur présence a été décrite sur les autres plasmides à bas nombre de copies (pour revue (Gerdes *et al.*, 2010; Salje, 2010).

La découverte et l'étude de différents systèmes de partition ont permis de réaliser une classification selon la protéine NTPase impliquées. Cette classification se décline en trois types. Le système de partition de type **I** rassemble les ATPase à motif "Walker box", le type **II**, les ATPase de type actine (communément nommée ParM) et le type **III**, les GTPase de type tubuline (nommé TubZ). Le système de partition de type **II**, présent sur le plasmide R1, est le mieux connu au niveau moléculaire. En revanche, le mécanisme moléculaire impliqué dans les systèmes de partition de type **I**, mis en évidence et étudié sur le plasmide F et le prophage P1 d'*E. coli*, n'est pas encore élucidé. Il est à noter que la plupart des chromosomes bactériens possède également un locus *par* et tous les systèmes de partition chromosomique identifiés sont de type **I** (Livny *et al.*, 2007).

Sur les chromosomes, le système ParABS est impliqué dans la ségrégation des origines bactériennes et serait également impliqué dans d'autres étapes de développement (Lee and Grossman, 2006) (**Introduction, Partie I.C**). Les plasmides possédant un système de partition de type **I** constituent donc des modèles d'étude intéressants pour la compréhension de la ségrégation active des chromosomes bactériens. Les systèmes de partition de type **III** ont été les derniers à être décrits et sont toujours en cours de caractérisation.

Des études sur la localisation des plasmides donnent un aperçu des événements impliqués dans leur ségrégation. Ces derniers présentent une localisation dynamique très ordonnée, renforçant l'idée d'un mécanisme actif impliqué dans la ségrégation des plasmides. Cependant, le positionnement exact dans la cellule est propre à chaque plasmide. Les plasmides F et P1 sont situés en position centrale lorsque les cellules sont de petite taille et aux positions $\frac{1}{4}$ et $\frac{3}{4}$ des cellules lorsqu'elles sont de grande taille (Niki and Hiraga, 1997; Gordon *et al.*, 1997). Dans le cas du plasmide R1, le positionnement est légèrement différent. Ils sont localisés au centre de la cellule ou aux pôles (Jensen and Gerdes, 1999). La présence des trois éléments des systèmes ParABS est requise et essentielle pour la localisation correcte des plasmides. L'absence d'un seul d'entre eux induit une distribution aléatoire des plasmides dans la cellule, conduisant à leur perte au cours des générations successives (Ogura and Hiraga,

1983). Ainsi, la réaction de partition est avant tout une réaction de positionnement des plasmides à l'intérieur de la cellule, permettant la répartition homogène des plasmides avant la division cellulaire. Les mécanismes moléculaires impliqués dans le processus de partition de chacun des systèmes types seront décrits dans la **partie II de l'introduction**.

2. Les systèmes de secours

Pour leur maintien, les plasmides sont soumis au contrôle de leur nombre de copies mais aussi de leur nombre d'unité de ségrégation. Lors de la duplication de l'ADN, des dimères ou éventuellement des multimères de plasmides peuvent apparaître suite à un évènement de recombinaison homologue entre les deux réplicons nouvellement synthétisés. La présence de dimères de plasmides réduit le nombre d'unité de ségrégation ce qui peut conduire à la perte du plasmide en fin de division. Des systèmes de résolution site spécifique permettent de convertir les dimères en monomères. Ces systèmes utilisent des recombinases à tyrosine, appelées XerCD/*xer* pour ColE1 (Summers and Sherratt, 1984), Cre/*loxP* pour P1 (Austin *et al.*, 1981) ou resD/*rfsF* pour le plasmide F (Lane *et al.*, 1986; Disqué-Kocher and Eichenlaub, 1993). Malgré tous les contrôles du maintien existant dans le cycle de vie du plasmide, il arrive parfois qu'un plasmide soit perdu au cours des générations successives. Un dernier système permet d'éliminer les bactéries dépourvues de plasmides et de ne maintenir dans la population bactérienne que les bactéries l'ayant hérité. Ce processus est effectué par des systèmes toxine-antitoxines dont le mode d'action, très bien décrit dans les revues suivantes, ne sera pas détaillé ici (pour revue (Gerdes *et al.*, 1997; Gerdes, 2000; Gerdes and Maisonneuve, 2012).

C. La ségrégation des chromosomes bactériens

Contrairement aux cellules eucaryotes, qui ont des phases temporellement distinctes de réplication, condensation des chromatides sœurs et ségrégation, les bactéries ont besoin d'organiser, de compacter et de séparer leurs chromosomes au fur et à mesure que les chromosomes frères sont générés.

Le développement des techniques de microscopie avec l'avènement de la microscopie à fluorescence a rendu possible des avancées importantes dans la compréhension de l'organisation du chromosome et du processus de ségrégation. Il est à présent possible, grâce à différentes types de sondes fluorescentes comme par exemple des protéines fluorescentes

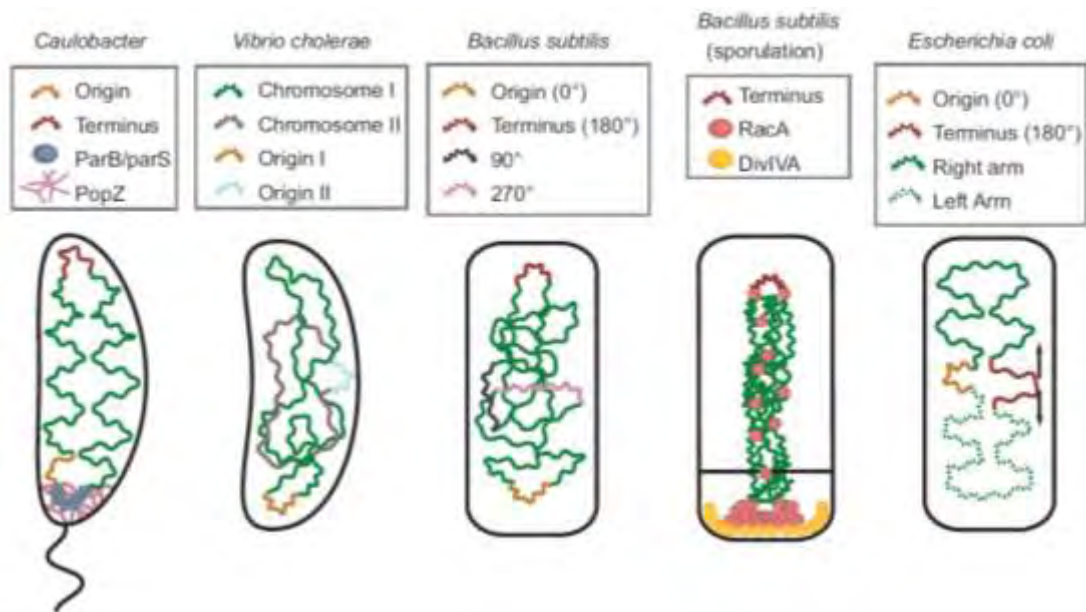


Figure 4 Localisation du chromosome bactérien dans la cellule

Schéma de l'organisation du chromosome dans plusieurs cellules bactériennes d'espèces différentes. Le chromosome (Chr) de *C. crescentus* est ordonné linéairement et ancré au pôle flagellé par le complexe *parS*/ParB/PopZ. Chez *V. cholerae* l'origine du ChrI est polaire, alors que celle de ChrII est centrale ; l'organisation de la masse des chromosomes est encore inconnue dans cette espèce. Chez *B. subtilis*, quatre loci ont été étudiés ; leur positionnement rappelle l'organisation chez *C. crescentus*, hors ancrage au pôle. Lors de la sporulation, cependant, on retrouve cet ancrage au pôle, auquel s'ajoute une compaction du chromosome. L'origine du chromosome d'*E. coli*, de même que son terminus, est centrale dans la cellule; les bras droit et gauche du chromosome sont ordonnés de part et d'autre du centre de la cellule.

Extrait de (Toro and Shapiro, 2010).

couplées à des protéines de fixation spécifique à l'ADN, de suivre au cours du cycle cellulaire, le devenir d'un locus chromosomique choisi (Gordon *et al.*, 1997; Wang *et al.*, 2008). Ces analyses, appelées FISH (fluorescence in situ hybridation) ou FROS (fluorescence repressor-operator system), ont permis de changer radicalement la vision que l'on avait de la bactérie. Celle-ci étant considérée à tort comme un "simple sac d'enzymes" où la majorité des processus se faisaient par diffusion passive, est au contraire hautement organisée (Thanbichler and Shapiro, 2008). Le chromosome avec son orientation précise et sa chorégraphie durant la ségrégation, ne fait pas exception à cette organisation. Ainsi la séparation de la plupart des chromosomes bactériens se décompose en trois étapes distinctes: la séparation et relocalisation rapide des origines nouvellement répliquées ; la ségrégation progressive des corps des chromosomes et la séparation des terminus de réplication avant la fermeture du septum de division. Le mode de ségrégation des chromosomes varie considérablement d'une espèce à l'autre. Chaque étape et mode de ségrégation seront discutés séparément en prenant pour exemple différents organismes modèles.

1. Ségrégation active des origines

Le mécanisme par lequel les origines nouvellement répliquées sont séparées a été l'objet de spéculations et d'investigations pendant plus d'un demi-siècle. Le modèle des réplicons proposé par (JACOB and BRENNER, 1963), a été le plus influant et a perduré pendant plus de trois décennies. Ce modèle postule que les deux origines nouvellement répliquées seraient attachées à l'enveloppe de la cellule et l'allongement des cellules par leur centre contribuerait à leur ségrégation. Il est maintenant clair que l'élongation des bactéries en forme de bâtonnets ne se limite pas à la croissance dans la zone médiane de la cellule, mais se produit sur des sites répartis dans toute la membrane (Typas *et al.*, 2011). De plus la vitesse de déplacement de l'origine, de l'ordre de 0.1-0.3 $\mu\text{m}/\text{min}$ (Webb *et al.*, 1998; Viollier *et al.*, 2004), excède la vitesse d'allongement cellulaire de 0.02 $\mu\text{m}/\text{min}$, ce qui va aussi à l'encontre de cette hypothèse.

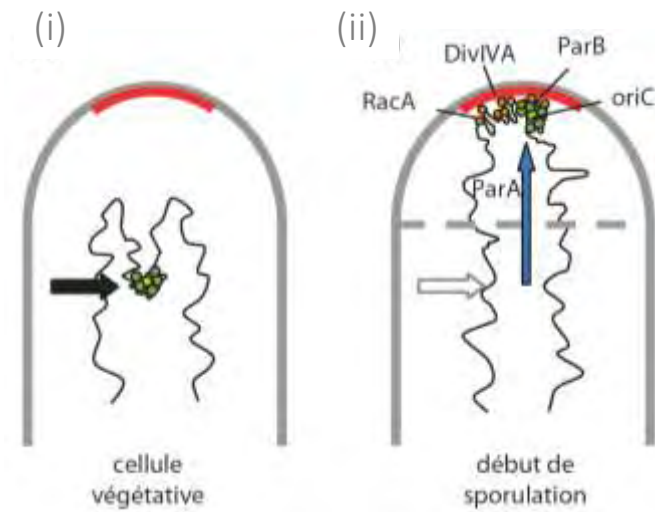
Des analyses de microscopie chez *E. coli* et *Bacillus subtilis*, ont révélé une localisation bien précise pour l'origine et le terminus de réplication dans des cellules nouvellement divisées (**Figure 4**). Le terminus est orienté vers l'un des pôles de la cellule et l'origine est positionnée au pôle opposé de la cellule (Webb *et al.*, 1997; Gordon *et al.*, 1997; Niki and Hiraga, 1998; Li

and Austin, 2002; Wang *et al.*, 2005). Après l'initiation de la réplication, les origines nouvellement répliquées migrent séparément aux positions $\frac{1}{4}$ et $\frac{3}{4}$ de la cellule alors que le terminus reste au centre de la cellule. La découverte des loci chromosomiques, homologues aux systèmes de partition actifs des plasmides à bas nombre de copies, a suggéré l'idée que ces systèmes pourrait être responsable du positionnement précis des origines dans la cellule (Ogura and Hiraga, 1983; Austin and Wierzbicki, 1983). En effet, la plupart des chromosomes bactériens portent des loci *parABS* chromosomiques (pour revue (Livny *et al.*, 2007), à l'exception des *Enterobacteriaceae*, avec comme chef de file la bactérie *E. coli* et des *Pasteurellaceae* (pour revue (Hiraga, 2000). Lorsqu'ils ont été testés, les systèmes ParABS chromosomiques se sont révélés capables d'assurer la ségrégation de plasmides dépourvus de leur propre système de partition (Lin and Grossman, 1998; Yamaichi and Niki, 2000; Godfrin-Estevenon *et al.*, 2002; Bartosik *et al.*, 2004; Dubarry *et al.*, 2006; Saint-Dic *et al.*, 2006). Leur localisation et plus particulièrement celle de leurs sites *parS* dans la région origine des chromosomes font de ces systèmes de bons candidats pour la ségrégation active de cette région.

Dans le cas de la bactérie *E. coli* dépourvue de systèmes ParABS, une séquence particulière a été identifiée comme étant importante pour la localisation des origines après ségrégation (Yamaichi and Niki, 2004). Cette séquence *migS* (*migration of séquence*) de 25pb est localisée à 200Kb de l'origine *OriC*. Elle serait la première région à être ségrégée du centre de la cellule aux positions $\frac{1}{4}$ et $\frac{3}{4}$, avant l'origine de réplication *OriC*. Néanmoins, la délétion de cette séquence n'entraîne pas de défaut de partition et son insertion dans un plasmide instable ne le stabilise pas (Yamaichi and Niki, 2004; Fekete and Chatteraj, 2005). De ce fait, *migS* n'est pas une séquence centromérique liée à un système de ségrégation Par et aucune protéine capable de reconnaître cette séquence n'a été mise en évidence depuis cette première étude. La localisation précise des origines à une position donnée dans des bactéries dépourvues de systèmes ParABS, a pour origine un mécanisme qui reste encore à décrire.

L'implication des systèmes ParABS chromosomiques dans le processus de ségrégation active des origines a pu être démontrée par l'apparition de deux phénotypes majeurs chez les bactéries étudiées : des défauts de positionnement de l'origine *OriC* et l'apparition de cellules anucléées. Ces deux phénotypes traduisent un défaut majeur de ségrégation. Chez *B. subtilis*, la perturbation des systèmes ParABS chromosomiques conduit à l'apparition de 1 à 2% de cellules

a



b

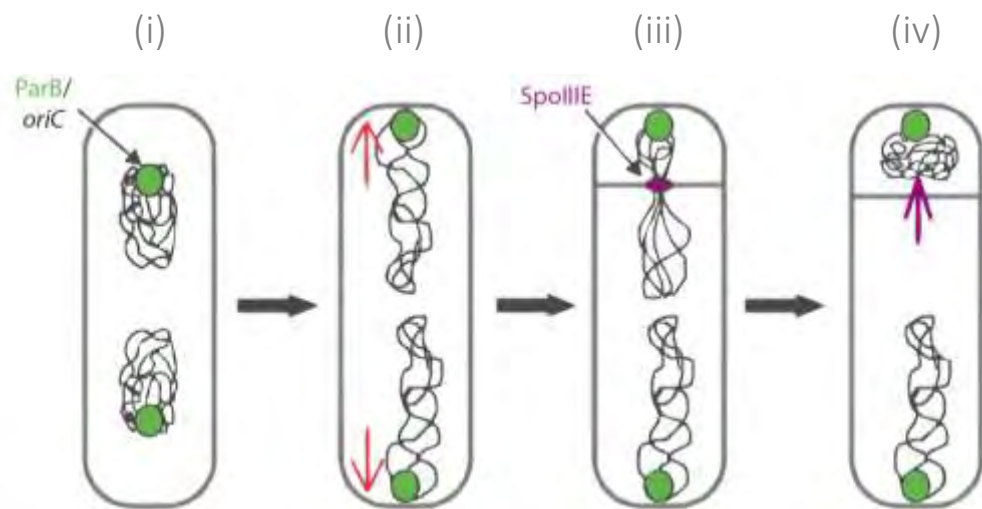


Figure 5 f Ségrégation lors de la sporulation de *B. subtilis*

-] **Ségrégation de la région origine.** (i) Cellule végétative. La flèche noire indique la position déterminée par le système parABS en croissance végétative. (ii) Lors de la sporulation, le couple RacA-DivIVA ancre la région origine au pôle, ParA aidant à amener l'origine à portée de DivIVA.
-] **Ségrégation de la masse du chromosome.** (i) Cellule végétative. (ii) En début de sporulation, le nucléoïde se réorganise en filament axial, d'un pôle à l'autre de la cellule. L'ancrage des origines aux pôles est représenté par les flèches rouges. (iii) Le septum de division asymétrique se forme sur le chromosome, et la translocase SpoIIIE se place au septum. (iv) SpoIIIE assure la translocation de 70% du chromosome de la cellule mère vers la préspore.
- ¶] Extrait de (Errington *et al.*, 2005).

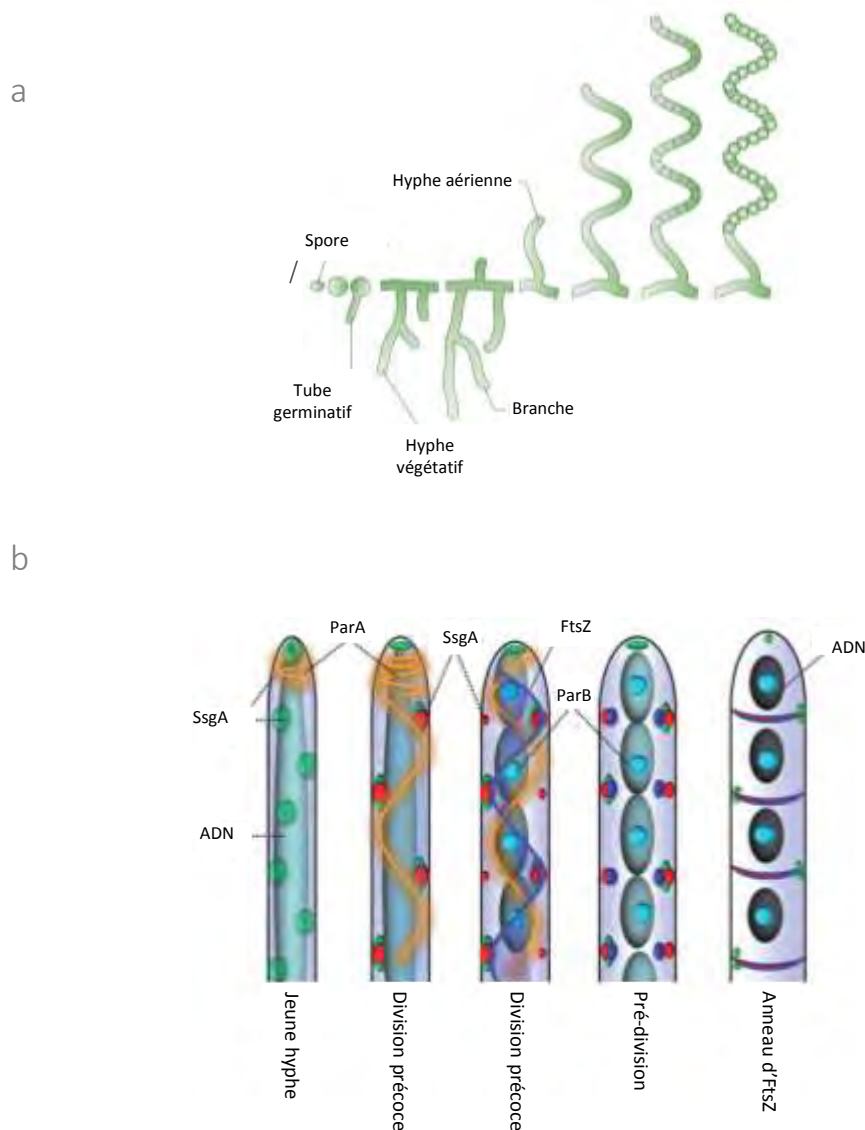
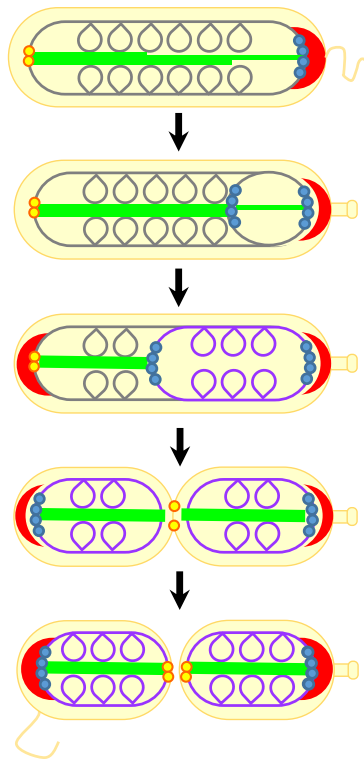


Figure 6 | Ségrégation et division lors de la sporulation de *S. coelicolor*

- a) **Modèle général de la sporulation.** La germination d'une spore implique le gonflement, la polarisation de la croissance et l'émergence d'un tube de germe qui se développe en hyphe. Les hyphes se développent, par extension, de la pointe et se ramifient en un mycélium végétatif qui pousse partout et en profondeur dans le substrat, donnant naissance à la colonie végétative. En réponse à l'épuisement des nutriments, les hyphes aériennes se différencient en une longue chaîne de compartiments, les prés pores. Extrait de (Flärdh and Buttner, 2009).
- b) **Modèle sur le contrôle de la formation de l'anneau d'FtsZ.** Chez les jeunes hyphes aériennes, les protéines SsgA-G sont des protéines impliquées dans la morphogénèse des hyphes et leur division. SsgA se localise suivant un modèle très dynamique alors que SsgB et FtsZ sont encore diffus. A ce moment-là, la protéine ParA est principalement localisée à la pointe des hyphes. Puis dans les premiers stades de division, SsgA et SsgB colocalisent temporairement soit d'un côté de l'hyphe aérienne, soit de l'autre, tandis que ParA forme un filament tout au long de l'hyphe. En présence de ParA, ParB s'assemble en complexe pour former des foci condensés sur l'ADN encore non condensé. Avant la division cellulaire spécifique à la sporulation, FtsZ s'assemble en forme de filaments en spirale le long de l'hyphe. Dans ces hyphes en pré-division, FtsZ et SsgB continuent à colocaliser et SsgB ancre FtsZ à la membrane. L'anneau FtsZ est alors formé à cette étape de sporulation. Il s'en suit la condensation des chromosomes, leur ségrégation, puis la formation des septums de divisions. Lorsque les spores sont matures, SsgA se localise de chaque côté des septums et marque les futurs sites de germination.
- c) Extrait de (Jakimowicz and van Wezel, 2012).

a



b

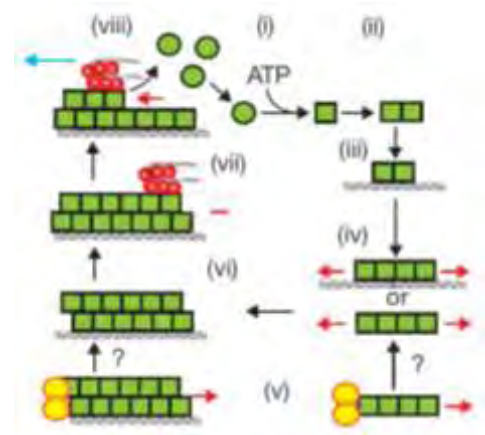


Figure 7 / Ségrégation du chromosome de *C. crescentus*

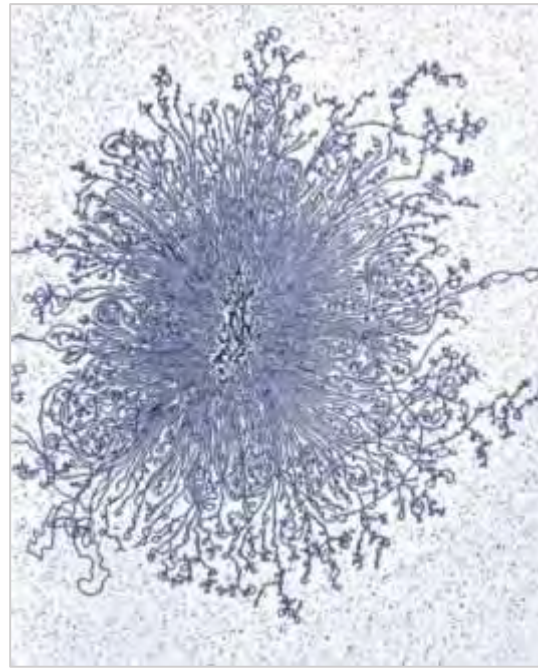
- a) **Schéma général de la ségrégation de la région origine.** Cette bactérie existe sous deux formes ; une forme pédonculée et une forme flagellée. Seule la cellule pédonculée peut entamer un cycle de division cellulaire, la cellule flagellée doit muer en cellule pédonculée pour le faire (pour revue (Brown *et al.*, 2009)). Au début du cycle cellulaire, la région origine est ancrée au pôle pédonculé par le complexe de partition ParB-*parS*, lui-même fixé au pôle par l'interaction de ParB (bleu) avec PopZ (rouge), une protéine polaire (Bowman *et al.*, 2008; Ebersbach *et al.*, 2008). ParA (vert) est présent au pôle opposé, ancré par la protéine polaire TipN (jaune) (Schofield *et al.* 2010), et s'étend dans la cellule sous forme de "comète". Lorsque la réplication de la région origine atteint le site *parS*, situé à 8Kb de *OriC*, la ségrégation du chromosome est initiée : deux complexes de partition sont formés puis l'un des deux contacte la "queue de la comète" de ParA qui se rétracte et entraîne le complexe de partition dans son sillage jusqu'au pôle opposé (Shebelut *et al.*, 2010), probablement selon un mécanisme de "pont brûlé" puisque ParB induit la dépolymérisation de ParA *in vitro* (Ptacin *et al.*, 2010). Le complexe de partition est ancré à ce pôle par PopZ (Bowman *et al.*, 2008)). Ainsi, une fois la division cellulaire achevée, chacune des cellules porte un complexe de partition au pôle le plus ancien (pédonculé ou flagellé). La structure de ParA se réorganise et TipN est recruté au site de division pour être en position pour le cycle suivant.
- b) **Mécanisme moléculaire impliquant le système ParABS dans la ségrégation.** (i) La protéine ParA (cercles verts clairs) fixe l'ATP, change de conformation (carrés verts foncé) et (ii) dimérise. L'homodimère ParA-ATP (iii) se fixe au nucléoïde et (iv) polymérise sur l'ADN ou librement (les flèches vertes indiquent la direction de polymérisation/dépolymérisation). (v) TipN (cercles jaunes) pourrait nucléer ou stabiliser le polymère de ParA au nouveau pôle. (vi) Les fibres de ParA s'assemblent en faisceaux. (vii) le complexe ParB-*parS* (cercles bleus) rencontre l'extrémité d'une fibre et s'y fixe. ParB stimule le détachement de ParA à l'extrémité des filaments et (viii) le complexe de partition avance avec la fin du filament de ParA qui se rétracte (la flèche rouge indique le mouvement du complexe de partition). Extrait et adapté de (Ptacin *et al.*, 2010).

anucléés dans un mutant $\Delta parB$ (nommé *spoOJ* chez cette bactérie), ce qui est 100 fois plus important que dans une souche sauvage (Ireton *et al.*, 1994). De plus la localisation des *OriC* est aussi perturbée pour des mutants $\Delta parA$ (nommé *soj* chez *B. subtilis*), $\Delta parB$ et $\Delta parAB$ (Lee and Grossman, 2006). L'implication de ces systèmes ParABS chromosomiques est aussi retrouvée pour d'autres bactéries comme *Pseudomonas putida* (Godfrin-Estevenon *et al.*, 2002), *Pseudomonas aeruginosa* (Lasocki *et al.*, 2007; Bartosik *et al.*, 2009), *Mycobacterium smegmatis* (Jakimowicz *et al.*, 2007), *Mycobacterium tuberculosis* (Maloney *et al.*, 2009) ou *Corynebacterium glutamicum* (Donovan *et al.*, 2010). De façon générale, à l'exception de *C. crescentus* (Mohl and Gober, 1997), la délétion des gènes codant les protéines ParA et ParB n'est pas létale pour les bactéries. Cependant, l'apparition d'une fréquence importante de cellules anucléées suggère que plus d'un système pourrait être responsable de la ségrégation de l'ensemble de la molécule d'ADN chromosomique.

Les systèmes ParABS semblent aussi être importants dans la ségrégation de l'origine du chromosome lors de la différenciation de certaines bactéries. Par exemple, lors de la sporulation de *B. subtilis*, la région origine est capturée au pôle de la cellule par l'intermédiaire du système ParABS et du système RacA-DivIVA, conduisant à son positionnement dans la préspore (Wu and Errington, 2003; Sullivan *et al.*, 2009) (**Figure 5**). Ainsi, une délétion de *parA/parB* (*soj-spoOJ*) entraîne un défaut d'orientation dans la préspore (Sharpe and Errington, 1996). Chez *Streptomyces coelicolor*, les systèmes ParABS semblent aussi être impliqués dans la sporulation. Lorsque la spore germe, elle forme un mycélium végétatif constitué d'hyphes polyploïdes. Le système ParABS aide au positionnement correct des chromosomes avant la septation de ces hyphes (Kim *et al.*, 2000; Jakimowicz and van Wezel, 2012) (**Figure 6**). Ainsi, dans des conditions particulières de croissance, comme la sporulation ou la synthèse d'hyphes, ces systèmes semblent importants pour la ségrégation et le positionnement des chromosomes.

Chez *Caulobacter crescentus*, le système ParABS est essentiel à la viabilité cellulaire et agit à deux niveaux : lors de la ségrégation des chromosomes en interagissant avec deux protéines PopZ et TipN ; et lors de la division en régulant la formation du septum de division (Mohl *et al.*, 2001) ; La ségrégation du chromosome chez *C. crescentus* est détaillée dans la **Figure 7**.

a



nucléoïde

b

*E. coli*

Figure 8 | Le nucléoïde bactérien

Les deux images sont mises à la même échelle pour imaginer le degré de compaction nécessaire au nucléoïde pour prendre place dans la bactérie.

- a) Visualisation en microscopie électronique à transmission du nucléoïde d'*E. coli* lié au cytochrome C, après fixation à l'acétate d'uranyl (Extrait de Designergenes Posters Ltd : en mémoire du docteur Ruth Kavenoff). La longueur d'une cellule d'*E. coli* mesure environ 1-3 μm alors que le nucléoïde étalé sur une grille de microscopie mesure 20 μm de longueur. En périphérie, le nucléoïde est décondensé et contient des sous domaines bouclés.
- b) Image d'une coupe d'*E. coli* en ME après cryofixation.

Il semble donc que l'importance variable des systèmes ParABS chromosomiques tient aussi de leur implication dans d'autres fonctions cellulaires et la redondance fonctionnelle avec d'autres systèmes impliqués dans la ségrégation du chromosome. En effet, les systèmes ParABS pourraient s'être adaptés aux différents contextes cellulaires en recrutant différentes protéines partenaires afin d'agir en synergie pour réaliser le mécanisme de ségrégation global.

2. Ségrégation du corps du chromosome

A l'inverse de la ségrégation active, rapide et dirigée des origines, la ségrégation du "corps" du chromosome est un processus lent et progressif qui suit la réplication (Nielsen *et al.*, 2006; Nielsen *et al.*, 2007) et ne nécessite probablement pas de mécanisme dédié. Plusieurs mécanismes non-spécifiques semblent toutefois agir en synergie ou non pour contribuer à la localisation précise de chaque région du chromosome dans la cellule (Toro and Shapiro, 2010) et au mécanisme global de ségrégation.

a) Organisation et compaction du chromosome - (les complexes SMC)

La taille du chromosome bactérien est une des contraintes majeures rencontrées lors du processus de ségrégation. Par exemple, le chromosome d'*E. coli* mesure environ 4.6 Mb ce qui correspond à un ruban d'ADN d'environ 2 mm de long. Sachant qu'une cellule d'*E. coli* mesure environ 1-3 μm , l'ADN doit nécessairement être compacté d'un facteur mille pour pouvoir être contenu dans la cellule sans pour autant interférer avec l'ensemble des processus cellulaires (**Figure 8**). Cette compaction du chromosome est encore mal connue mais se fait de manière organisée jusqu'à la formation d'une masse nucléoprotéique, appelée nucléoïde. Selon les conditions de croissance, le nucléoïde peut prendre l'aspect d'un O, d'un C, d'une forme bilobée ou toroïdale en microscopie (Zimmerman, 2006). Ce changement d'aspect du nucléoïde suggère un état dynamique de l'ADN qui est sans cesse soumis à des forces de compaction et d'expansion suivant l'état physiologique de la cellule (état transcriptionnel de la cellule et/ou le passage de la fourche réplivative). Watson et Crick ont publié en 1953 "Bien qu'il soit difficile pour le moment de voir comment tous ces processus se produisent sans s'emmêler, nous ne pensons pas que ce soit insurmontable". Dans cette partie, nous décrirons comment la condensation très organisée du chromosome pourrait participer à la séparation des chromatides sœurs en rythme avec la réplication.

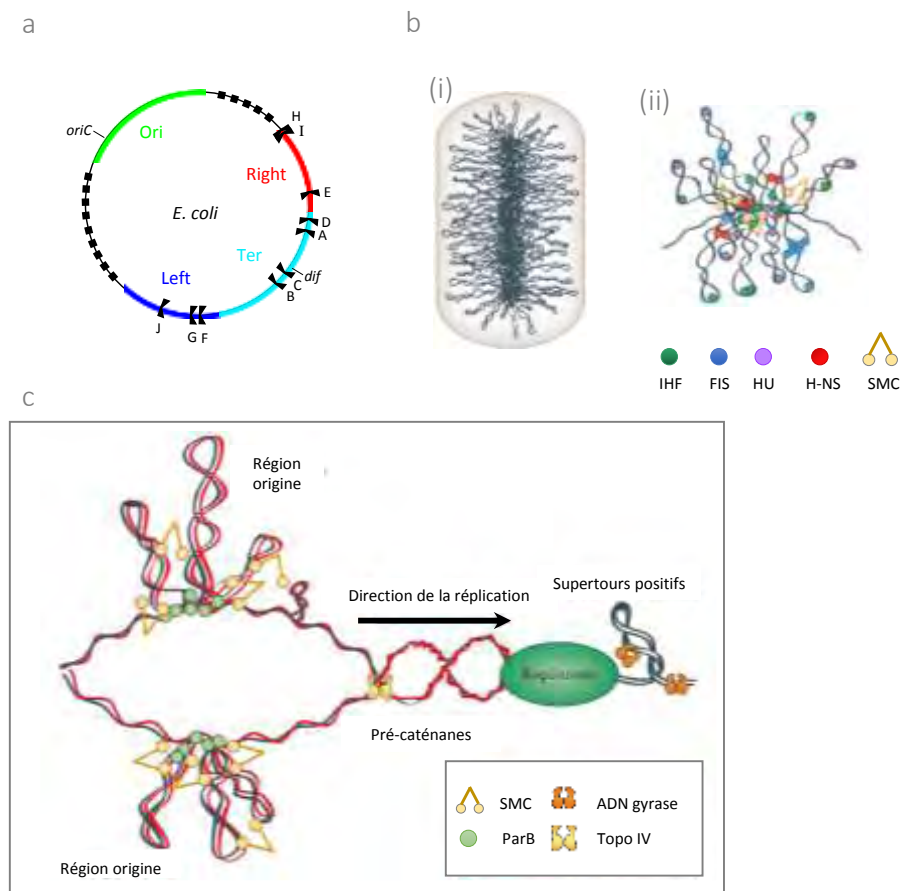


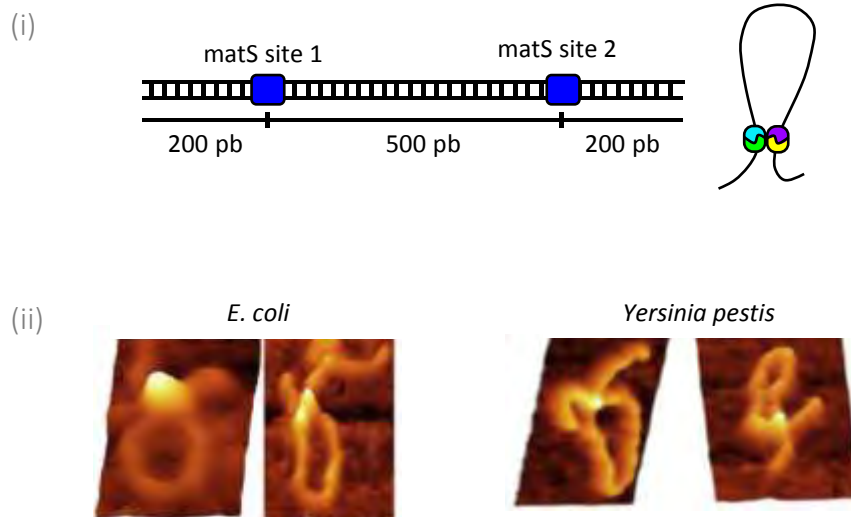
Figure 9 | Organisation et compaction du intra-chromosomique du chromosome bactérien

- a) **Organisation en macrodomaines.** Représentation schématique des macrodomaines et des régions non-structurées du chromosome d'*E. coli*. Les macrodomaines sont des régions à l'intérieur desquelles les contacts entre les loci sont plus fréquents qu'avec l'extérieur de ces domaines. Le cercle représente la carte génétique du chromosome. Les zones colorées représentent les différents macrodomaines, les bandes noires, les régions non structurées. Sont aussi indiqués les dix sites Ter (arrêt des fourches de réplication, A à J), l'origine de réplication *OriC*, le centromère putatif *migS*, et le site *dif* de résolution des dimères. Extrait de (Toro and Shapiro, 2010).
- b) **Compaction intra-chromosomique.** (i) Représentation schématique des boucles d'ADN superenroulées autour du noyau dense. Les domaines topologiquement isolés sont en moyenne de 10 Kb et sont donc susceptibles d'englober plusieurs boucles plectonémiques ramifiées. (ii) Représentation schématique des différentes petites protéines associées au nucléoïde ainsi que les protéines du complexe SMC. Ces protéines induisent des coudes au niveau de l'ADN et servent également à rapprocher des loci chromosomiques par pontage. Extrait de (Reyes-Lamothe *et al.*, 2012).
- c) **Réplication et compaction.** Le schéma illustre la progression de la fourche de réplication et la compaction concomitante de la zone d'origine. La réplication génère des supertours positifs avant la fourche de réplication. Ces derniers peuvent diffuser en aval du réplisome, produisant ainsi des pré-caténanes. Les supertours positifs sont éliminés par les topoisomérases de type II comme la gyrase et la topoisomérase IV (Topo IV). Les caténanes sont dissociées par la Topo IV. Les régions d'origine nouvellement répliquées sont surement compactées par le complexe SMC qui est recruté à l'origine par ParB et par l'action des petites protéines associées au nucléoïde. Extrait de (Reyes-Lamothe *et al.*, 2012).

Le surenroulement négatif joue un rôle premier dans le mécanisme de compaction et de ségrégation en repliant l'ADN sur lui-même et en l'éloignant de manière indirecte de l'ADN directement contigus (comme l'ADN nouvellement répliqué). Ce surenroulement est crucial pour la survie de la bactérie car il contrôle l'activité de certains promoteurs et le métabolisme en général (Holmes and Cozzarelli, 2000). Plusieurs types de protéines associées à l'ADN agissent dans le maintien et les modifications de sa condensation et de sa topologie. Par exemple, les topoisomérases de type **II**, comme la gyrase, maintiennent la super-hélicité négative de l'ADN en enlevant les supertours positifs générés au cours de la réplication et la transcription (Wang, 2002). Ce surenroulement négatif de l'ADN double brin, génère des boucles plectonémiques (**Figure 9**) à l'origine de domaines super-enroulés de 10 Kb environ et topologiquement isolés les uns des autres (Worcel and Burgi, 1972; Postow, 2004). Ces domaines topologiques protègent le chromosome de la relaxation de l'ADN et aide à la réparation des cassures double-brin en maintenant de façon cohésive les extrémités cassées (Postow, 2004). Le surenroulement négatif ne peut à lui seul expliquer le degré de compaction du chromosome. La moitié du chromosome semble être contrainte par des petites molécules basiques non essentielles, les NAPs (Protéines Associées au Nucléoïde) qui s'associent à l'ADN et le condensent. Par exemple chez *E. coli*, il existe une dizaine de NAPs différentes dont HU, H-NS, IHF et FIS. Les NAPs ne sont pas des homologues d'histones et de ce fait, ne forment pas de nucléosome. Leur propriété commune est leur capacité à lier l'ADN et à le courber différemment selon la NAP fixée (Wang *et al.*, 2013). Certaines participent à la compaction locale du chromosome en créant des coudes dans l'ADN (IHF, HU) ou d'autres pourraient stabiliser des boucles adjacentes par pontage (H-NS, FIS) (Dame *et al.*, 2005; Dillon and Dorman, 2010; Wang *et al.*, 2011).

En plus d'être organisé localement, le chromosome bactérien est structuré en macrodomaine (Niki *et al.*, 2000). Ces grandes régions (800 Kb-1 Mb) ont été identifiées chez *E. coli* mais pourraient être une caractéristique commune à de nombreux chromosomes bactériens. Quatre macrodomaines et deux régions non structurées ont été recensés et sont décrits en **Figure 9a**. Les macrodomaines ont été définis selon deux propriétés : leur localisation intracellulaire précise et spécifique en opposition au comportement dynamique des régions non structurées, et leur faible fréquence de rencontre au cours du cycle cellulaire (Niki *et al.*, 2000; Valens *et al.*, 2004). Ce niveau d'organisation suggère l'existence de partenaires dédiés

a



b

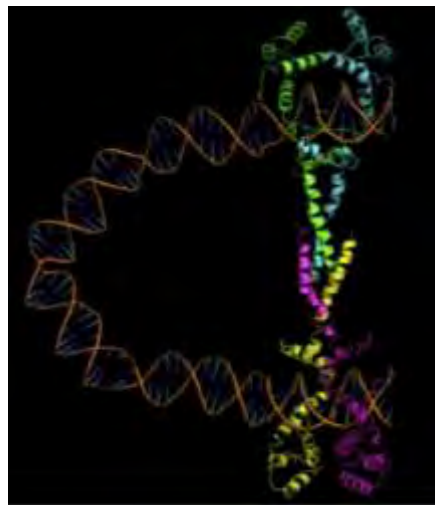


Figure 10 *Visualisation des boucles d'ADN formées par MatP*

- a) **Images d'AFM de boucles d'ADN formées par MatP.** L'organisation de l'ADN utilisée dans l'analyse d'AFM est montrée en (i). Les images d'AFM ont été réalisées avec MatP d'*E. coli* et MatP d'*Y. pestis* liés à l'ADN (i), contenant les deux sites *matS* (ii). Ces images montrent que la liaison de deux dimères de MatP forme une boucle de 500 pb à l'image du petit schéma en (i) à droite.
- b) **Modèle moléculaire d'une boucle d'ADN formée par un tétramère de MatP d'*E. coli*.** Un dimère de MatP est représenté en bleu et vert, et l'autre en jaune et violet. Chacun des deux dimères de MatP fixé à leur site *matS* vont interagir pour former un tétramère. Ce dernier permet le rapprochement des deux sites *matS* et la création d'une boucle d'ADN.
- Adaptée de (Dupaigne *et al.*, 2012).

à la structuration de chaque domaine. En effet, des analyses bio-informatiques ont identifié 21 séquences *matS* dans le macrodomaine *ter* (Mercier *et al.*, 2008). Ces séquences sont reconnues spécifiquement par la protéine MatP. En l'absence de cette protéine, on observe une augmentation de la mobilité du macrodomaine *ter*, ainsi que de la fréquence de recombinaison entre domaines adjacents. Ces données suggèrent un rôle d'organisateur de la région *ter* pour la protéine MatP. De plus, des analyses structurales ont récemment montré que deux dimères MatP fixés sur deux sites *matS* distants, pouvaient s'oligomériser en tétramère et conduire au rapprochement de deux sites (Dupaigne *et al.*, 2012) (**Figure 10**). L'analyse *in vivo* de mutants incapables de tétramériser confirme le rôle de MatP dans la condensation du macrodomaine *ter*. Comme il a été dit plus haut, la protéine H-NS est aussi capable de former des ponts protéiques entre molécules d'ADN dans les 4 macrodomaines et les deux régions non structurées (Dame *et al.*, 2005; Wang *et al.*, 2011). MatP pourrait être un exemple de condensine spécifique du macrodomaine *ter* alors qu'H-NS serait une condensine universelle du chromosome. Cette organisation d'ordre supérieur qui définit un macrodomaine ne semble pas jouer de rôle central dans le processus de ségrégation des chromosomes. Cependant, l'implication des complexes H-NS dans la régulation de gènes dont les régions promotrices sont sensibles à la topologie de l'ADN (Hommais *et al.* 2001 ; Wagner *et al.* 2000) ouvre d'autres perspectives quant aux diverses fonctions de ces protéines. Un autre acteur majeur de la compaction du chromosome est le système SMC (Structural Maintenance of Chromosome) (**Figure 9 et Tableau 1**).

| Organisme | Taille du chromosome (Mb) | <i>par</i> locus | SMC ^a | ADN translocase |
|-------------------------------|---------------------------|---|------------------|---------------------------|
| <i>Escherichia coli</i> | 4,6 | / | MukB (MukFE) | FtsK |
| <i>Bacillus subtilis</i> | 4,2 | ParA(Soj) B (SpoOI) / <i>parS</i> | BsSMC (ScpAB) | SfrA/SpoIIIE (vegetative) |
| | | | | SpoIIIE (sporulation) |
| <i>Caulobacter crescentus</i> | 4 | ParAB/ <i>parS</i> | SMC (ScpAB) | FtsK |
| <i>Vibrio cholerae</i> | chrI: 3,0 | ParA ₁ B ₁ / <i>parS</i> ₁ | MukB (MukFE) | FtsK |
| | chrI: 1,0 | ParA ₂ B ₂ / <i>parS</i> ₂ | / | utilise FtsK du ChrI |

^a Les protéines accessoires des protéines SMC sont indiquées entre parenthèses

/ pas de données

Adapté de Reyes et Lamothe, 2012

Tableau 1 | Acteurs impliqués dans la ségrégation des chromosomes bactériens

Ce système a initialement été identifié chez les eucaryotes où il est impliqué dans de nombreux processus comme la cohésion des chromatides sœurs, la condensation des chromosomes avant la mitose et la recombinaison (Nasmyth and Haering, 2005). Des

homologues structuraux de ce complexe ont été identifiés chez *B. subtilis* et *E. coli*, nommés SMC_{BS} ou MukB respectivement (Holmes and Cozzarelli, 2000). En absence ou en surexpression de cette protéine, le chromosome sera partiellement décondensé ou hypercondensé, démontrant son lien direct dans la compaction du chromosome (Hiraga *et al.*, 1989; Britton *et al.*, 1998; Volkov *et al.*, 2003). Cependant son rôle ne se restreint pas à la condensation du chromosome. L'apparition de cellules anucléées, de chromosomes guillotisés par le septum de division dans des mutants Δ SMC (Hiraga *et al.*, 1989; Britton *et al.*, 1998) et la localisation de MukB et de SMC_{BS} au niveau des origines de réplication (Danilova *et al.*, 2007; Sullivan *et al.*, 2009), semblent indiquer un rôle de ces complexes dans la ségrégation. Chez *B. subtilis* et chez *Streptococcus pneumoniae*, SMC interagit avec la région origine au niveau du complexe ParB/*parS* (Sullivan *et al.*, 2009; Gruber and Errington, 2009; Minnen *et al.*, 2011). Ainsi le système SMC pourrait initier dans cette région la compaction des chromosomes frères nouvellement répliqués et participer à la ségrégation du chromosome.

b) L'expulsion par le réplisome et l'implication de la transcription

Les machines cellulaires de réplication, transcription, traduction co-transcriptionnelle et de translocation des protéines membranaires pourraient participer à la ségrégation des chromosomes en apportant la force motrice à ce processus. Ce modèle tient de la découverte chez *B. subtilis* du positionnement de ces machineries au centre de la cellule (Lin and Grossman, 1998; Lemon and Grossman, 2001). Par exemple, l'ARN polymérase, immobilisée au centre de la cellule pourrait tracter l'ADN au fur et à mesure de la synthèse de l'ARN messager (Gelles and Landick, 1998; Wang, 1998; Kruse *et al.*, 2006). En accord avec cette hypothèse, il a été montré que l'inhibition de la transcription empêche la ségrégation correcte de l'ADN chez *B. subtilis* (Dworkin and Losick, 2002) et chez *E. coli* (Kruse *et al.*, 2006). De même, la machinerie de réplication appelée réplisome, pourrait diriger l'ADN nouvellement répliqué vers les deux côtés de la cellule au fur et à mesure de sa réplication et conduire à sa séparation. Depuis, l'avancée des connaissances sur la dynamique et la physique de la machinerie de réplication amène à penser que ces mécanismes n'ont que peu d'effet, voire aucun. L'amélioration des techniques de microscopie a par exemple révélé que le réplisome n'est finalement pas si stable, et oscille autour d'une position moyenne (Reyes-Lamothe *et al.*, 2008). De plus, les études des propriétés physiques du chromosome ont révélé que celui-ci est assez souple et élastique et la force d'expulsion du réplisome serait donc rapidement absorbée par les déformations de la

molécule (Dworkin and Losick, 2002). Par ailleurs, la nécessité de “démêler” les liens d’intercaténation entre les chromosomes frères après leur répllication empêche probablement leur expulsion immédiate vers les extrémités opposées de la cellule.

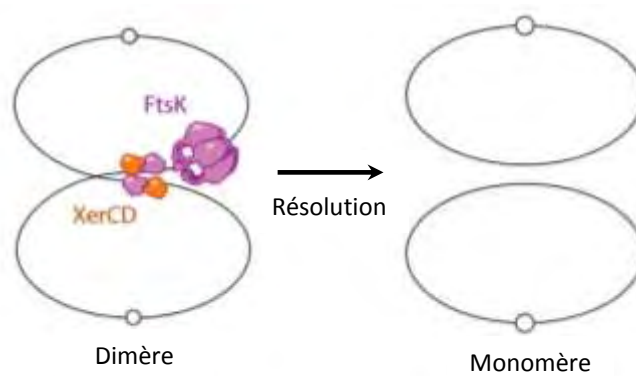
Néanmoins, tous ces systèmes pourraient servir de point de départ pour orienter les segments d’ADN nouvellement répliqués à la sortie du réplisome et cette orientation pourrait avoir un impact non négligeable sur l’efficacité des autres acteurs de la ségrégation.

c) L’entropie

Jun and Mulder, 2006, ont proposé que les forces d’entropie puissent être suffisantes pour réaliser la séparation spontanée de chromosomes frères résidant dans un même espace cylindrique confiné (Jun and Mulder, 2006; Jun and Wright, 2010). L’évolution a conduit à la complexification des processus biologiques par l’utilisation de macromolécules afin d’augmenter l’efficacité des réactions. La séparation par l’entropie contribue probablement à la ségrégation des chromosomes mais il semble peu probable que ce processus agisse seul. Une récente étude couplant analyse microscopique et modélisation *in silico*, a permis d’approfondir la compréhension de l’influence de l’entropie sur la ségrégation bactérienne (Junier *et al.*, 2013). A elle seule, l’entropie ne parviendrait pas à réaliser la séparation spontanée des chromosomes frères ni à expliquer la chorégraphie précise des chromosomes lors du cycle cellulaire. Par contre, le couplage de l’entropie avec le positionnement précis de certains loci chromosomiques (région Origine et/ou Terminus) dans la cellule et l’organisation du chromosome en macrodomaines pourrait générer une force suffisante à la séparation des chromosomes (Junier *et al.*, 2013).

Aucun mécanisme actif n’est spécialement dédié au mécanisme de ségrégation du corps du chromosome. La séparation des deux chromosomes nouvellement synthétisés semble être le résultat d’un ensemble de mécanismes individuels non spécialisés dans la ségrégation. L’entropie, l’orientation de l’ADN en sortie du réplisome et tous les acteurs de la condensation des chromosomes semblent agir en synergie pour conduire progressivement à la répartition passive des chromosomes frères dans chaque cellule. L’apparition de cellules anucléés ou de défauts de ségrégation dans des mutants présentant des défauts de surenroulement ou délétés pour un des composants du système SMC, même en faible pourcentage, va dans le sens de cette hypothèse (Niki *et al.*, 1991; Painbeni *et al.*, 1997; Britton *et al.*, 1998; Jensen and Shapiro,

a



b

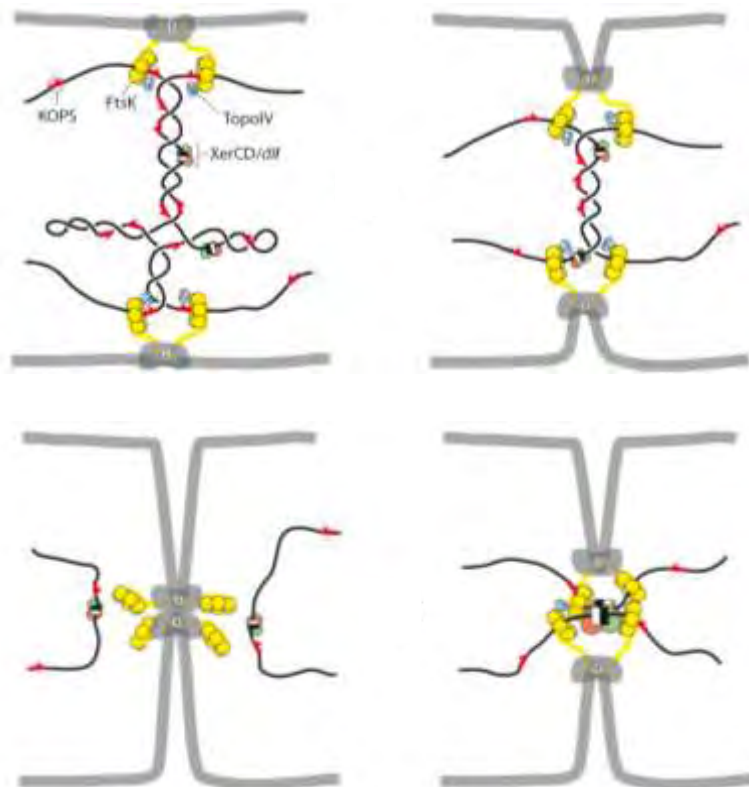


Figure 11 *FtsK, l'acteur clé de la fin de la ségrégation chez E. coli*

- a) **Résolution des dimères.** La recombinaison induite par le système de réparation lors de la réplication peut entraîner des cross-overs, générant des dimères de chromosomes, qui sont convertis en monomères par les recombinases XerCD, agissant au site *dif* dans la région *ter*. Extrait de (Reyes-Lamothe *et al.*, 2012).
- b) **Actions de FtsK.** FtsK (en jaune) est localisé au septum de division et y recrute la Topoisomérase IV (en bleu). L'activité translocase de FtsK, guidée par sa lecture des KOPS, permet de dégager l'ADN du septum de division. L'action de la TopoIV, stimulée par FtsK, démêle les chromosomes. Dans le cas d'un dimère de chromosomes, les deux copies des sites *dif* se retrouvent piégées au niveau du septum. FtsK active alors les recombinases XerCD spécifiques du site *dif* afin de résoudre le dimère. Figure empruntée à F. Cornet.

1999; Sullivan *et al.*, 2009; Schwartz and Shapiro, 2011). De plus, la modélisation mathématique de deux anneaux en polymère flexible suggère que la ségrégation ne pourrait nécessiter aucun autre mécanisme particulier que le démêlage par la Topoisomérase IV, des liens d'intercaténation des molécules d'ADN nouvellement synthétisées. Ainsi, tous les mécanismes décrits ci-dessus semblent agir en soutien du mécanisme de séparation passif (Marko, 2009).

3. La ségrégation du terminus : FtsK/SpoIII E

La dernière étape de partition correspond à la duplication et à la séparation des terminus au niveau du septum de division. L'acteur majeur de cette étape est une ADN translocase présente au septum de division (appelée FtsK chez *E. coli* et *C. crescentus*, et SpoIII E chez *B. subtilis*) (Wu *et al.*, 1995; Bigot *et al.*, 2007). Chez *E. coli*, FtsK, est capable de reconnaître des séquences KOPS de huit paires de bases (Bigot *et al.*, 2007). En suivant l'orientation biaisée de ces séquences, FtsK tracte le chromosome jusqu'au site *dif* ce qui permet de séparer les deux régions terminus et de dégager le septum de division avant la septation. Pour éviter la guillotine des chromosomes à cette étape, tous les liens entre chromosomes frères, doivent être rompus. Les liens d'intercaténation sont résolus par la Topoisomérase IV. Par ailleurs, la réplication génère dans 15% des cas des dimères de chromosome (Steiner and Kuempel, 1998). Un mécanisme de recombinaison site-spécifique (appelé XerCD/*dif* chez *E. coli* ou RipX et CodV chez *B. subtilis*), est dédié à cette fonction chez les bactéries (Lemon and Grossman, 2001; Lesterlin *et al.*, 2004; Kaimer *et al.*, 2011; Kaimer and Graumann, 2011) ainsi que chez les archées (Cortez *et al.*, 2010)) (**Figure 11**). La recombinase XerCD est régulée et recrutée au septum de division par l'ADN translocase, FtsK (Bigot *et al.*, 2007). FtsK est donc directement impliquée dans le mécanisme de résolution des dimères chez *E. coli*, alors que chez *B. subtilis*, cette translocase, appelée SpoIII E, serait uniquement requise lors de la sporulation (**Figure 5**). Par un mécanisme similaire à FtsK de reconnaissance de motifs biaisés, SpoIII E permet la translocation de 75% du chromosome dans la préspore, (SRS, SpoIIA Recognition Sequence), (Wu and Errington, 1994; Wu *et al.*, 1995; Wu *et al.*, 1998; Ptacin *et al.*, 2008; Sullivan *et al.*, 2009).

Les dimères de chromosome sont des problèmes spécifiques aux chromosomes circulaires et les linéaires en sont exemptés. Or, FtsK et Xer sont présents chez la plupart des bactéries ayant des chromosomes circulaires, mais aussi chez *S. coelicolor* et *Borrelia*, qui ont

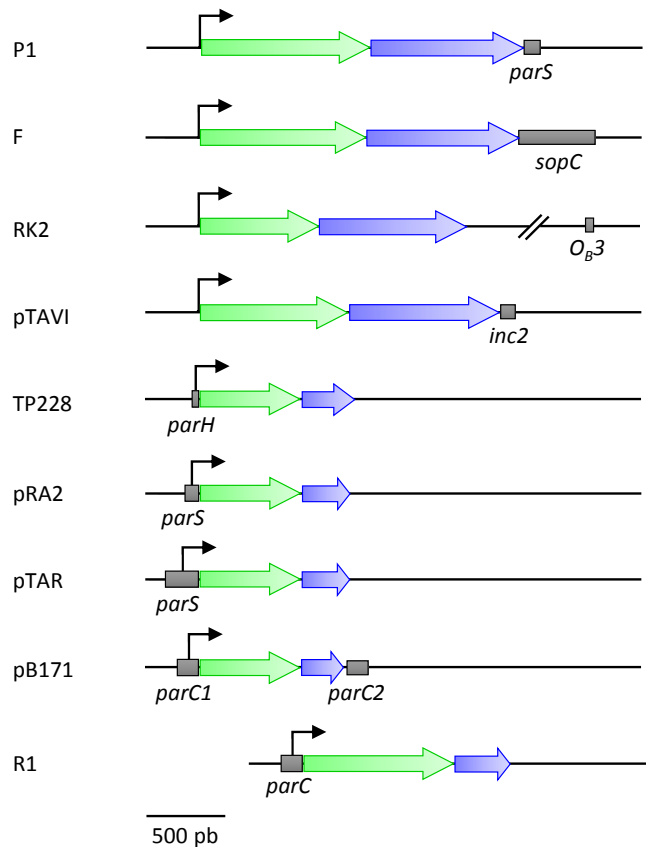
des chromosomes linéaires (Flärdh and Buttner, 2009; Chaconas and Kobryn, 2010). L'utilité de ces systèmes ne semble pas être restreinte à leur rôle dans la résolution des dimères. Par exemple, chez *S. coelicolor*, on observe un taux de perte élevé des extrémités des chromosomes linéaires lorsque l'activité de FtsK est réduite (Chen *et al.*, 2002; Wang *et al.*, 2007; Ausmees *et al.*, 2007). Ainsi, FtsK est important pour dégager l'ADN du septum de division lors de la septation des hyphes polyploïdes. Une récente étude attribue un rôle à FtsK dans la ségrégation active et la séparation des régions terminus de deux chromosomes frères monomériques (Stouf *et al.*, 2013). Par l'utilisation d'un système de marquages bicolores, la ségrégation de deux loci distincts sur chacun des chromosomes frères a été suivie. Dans plus de 99% des cas, le locus *ter* est séparé en dernier donc après les loci situés en dehors du locus *ter*. De même, dans 94% des cas, le site *dif* est séparé après les loci qui lui sont éloignés de 100 Kb, suggérant ainsi l'existence d'une ségrégation ordonnée de la région entourant le site *dif* (200 Kb). Ni le système XerCD / *dif*, impliqué dans la résolution des dimères, ni RecA, responsable de la formation de dimères, ne sont nécessaires pour la séparation ordonnée de la région *dif*. L'utilisation de différents mutants FtsK, y compris un allèle déficient pour la reconnaissance des sites KOPS, a permis de montrer l'importance de FtsK dans la ségrégation orientée de la région *ter*, bien que cette activité ne soit essentielle que dans le cas des dimères de chromosome. La protéine MatP (décrite plus haut), semblerait aussi être impliquée dans ce processus orienté. Cependant, le rôle exact des interconnexions entre le divisome et les systèmes d'organisation (MatP/*matS*) et de ségrégation (FtsK/KOPS) des régions *ter*, sont toujours en cours d'investigation.

La partition des chromosomes bactériens est un mécanisme complexe qui reste encore aujourd'hui, très largement incomprise d'un point de vue moléculaire. Plusieurs partenaires interviennent tant sur l'organisation que sur la dynamique des chromosomes bactériens pour permettre le processus de séparation global. En effet, les trois systèmes, ParABS, SMC et la translocase FtsK, semblent être directement impliqués dans la ségrégation des chromosomes bactériens. Cependant, dans la plupart des cas, aucun de ces systèmes n'est strictement nécessaire ou suffisant pour la ségrégation (**Tableau 1**). FtsK agit préférentiellement au niveau de la séparation des termini de réplication, alors que les systèmes SMC et ParABS agissent préférentiellement dans l'organisation du chromosome, le positionnement et la facilitation de la ségrégation des origines nouvellement répliquées. Si pour les chromosomes bactériens, le

Figure 12 J Diversité des organisations génétiques des opérons par plasmidiques

Les flèches vertes représentent les gènes codant les NTPase, les flèches bleues, les gènes codant les protéines de liaison au centromère. La localisation des centromères est indiquée par les rectangles gris. Le module de partition R1, a une organisation génétique similaire, mais il est phylogénétiquement distinct.

Figure adaptée de (Hayes and Barillà, 2006).



| Type | Caractéristique générales | plasmide/ espèce | Nomenclature des protéines NTPase | Protéine de fixation au centromère | | Centromère |
|------|--|-----------------------------------|-----------------------------------|------------------------------------|------------------|----------------|
| | | | | Nomenclature | (motif impliqué) | |
| I | ATPase à motif walker box | | | | | |
| Ia | Régulation de l'opéron <i>parAB</i> par fixation de ParA (HTH), grande protéine ParA et ParB | P1, P7 | ParA | ParB | HTH | <i>parS</i> |
| | | F, N15 | SopA | SopB | | <i>sopC</i> |
| | | RK2, RP4 | IncC1/2 | KorB | | <i>OB1-10</i> |
| Ib | Régulation de l'opéron <i>parAB</i> par ParB, ATPase plus courte que ParA la (pas de motif HTH) | pB171 (<i>par</i> ₂) | ParA | ParB | RHH | <i>parC</i> |
| | | TP228 | ParF | ParG | | <i>parH</i> |
| | | pSM19035 | δ2 | ω2 | | <i>pars1-3</i> |
| | | pTAR | ParA | ParB | | <i>parS</i> |
| Ic | Système chromosomique, ATPase plus courte que ParA la (pas de motif HTH) | Bsu, Tth... | Soj | SpoOJ | HTH | <i>parS</i> |
| | | Ccre, Ppu, Pae, Scoe... | ParA | ParB | | <i>parS</i> |
| II | ATPase de type actine | pB171 (<i>par</i> ₁) | ParM | ParR | RHH | <i>parC</i> |
| | | R1 | ParM | ParR | | <i>parC</i> |
| | | pLS32 | AlfA | AlfB | | <i>parN</i> |
| III | GTPase de type tubuline | pBtoxis | TubZ | TubR (wHTH) | wHTH | <i>tubC</i> |
| | | pXO1 | RepX | | | <i>tubC</i> |
| IV | Une seule protéine à domaine HTH (fixation au centromère) et motif coiled coil (pour la multimérisation) | pSK1 | par | par | n.i. | <i>DR/IR</i> |

Tableau 2 J Diversité des nomenclatures et principales caractéristiques des systèmes de ségrégation

Exemples de la nomenclature des systèmes actifs de partition. Pour pB171 : *par*₁ et *par*₂ sont deux systèmes de partition actifs portés par ce plasmide. Les abréviations d'espèces pour le type Ic sont les suivantes : Bsu *Bacillus subtilis*, Tth *Thermus thermophilus*, Ccre *Caulobacter crescentus*, Ppu *Pseudomonas putida*, Pae *Pseudomonas aeruginosa*, Scoe *Streptomyces coelicolor*. Pour le plasmide pXO1 (type III) le symbole Ø indique que ce système ne semble pas comporter de protéine de fixation à l'ADN autre que la GTPase. n.i.: non identifié.

rôle des systèmes ParABS n'est pas encore pleinement défini, pour les plasmides à bas nombre de copies les systèmes de partition sont absolument essentiels à leur maintien au cours des générations cellulaires. Leur étude représente un grand intérêt pour la compréhension de la ségrégation des réplicons bactériens et l'état actuel des connaissances sur les mécanismes moléculaires de ces systèmes, sera détaillé dans la **partie II de l'introduction**.

II. Mécanismes moléculaires des systèmes de partition

La partition de la grande majorité des plasmides à bas nombre de copies connus est gérée par un module à trois composantes codé par un locus *par*. Ce locus est composé d'un segment d'ADN appelé "séquence centromérique" et de deux gènes codant respectivement pour une protéine NTPase et une protéine qui se fixe spécifiquement sur la séquence centromérique. La transcription de cet opéron est directement autorégulée par les protéines Par qu'il code (Friedman and Austin, 1988; Hayes *et al.*, 1994; Jensen *et al.*, 1994; Hirano *et al.*, 1998). L'organisation génétique générale de ce locus, décrite en (**Figure 12**), peut être différente selon le type de plasmide. Les composants de ce système sont génériquement appelés, ParA, ParB et *parS* pour la séquence centromérique. Cependant, il existe une grande hétérogénéité de nomenclature entre les différents systèmes de partition étudiés (**Tableau 2**). La protéine ParA, considérée comme la protéine motrice du système, est une NTPase capable de positionner précisément les plasmides à l'intérieur de la cellule. Ce positionnement permet d'assurer la répartition équitable des plasmides dans chaque cellule. La protéine ParB, l'autre acteur de la partition, est une protéine adaptatrice entre la molécule d'ADN et la protéine motrice. Au vue de la grande diversité existant pour les protéines ParB et les sites *parS*, c'est en général la recherche d'un homologue de ParA, partenaire le mieux conservé, qui permet l'indentification *in silico* des systèmes ParABS. ParA est à l'origine de la classification des différents systèmes de partition. Il existe trois classes majeures selon le type de NTPase impliquée.

Le système de partition de type **I** rassemble les ATPases à motif "Walker box ", le type **II**, les ATPase de type actine et le type **III**, les GTPase de type tubuline. Le modèle général de la partition implique une fixation des protéines de liaison sur leurs sites centromériques après réplication du plasmide. Bien qu'il n'existe que peu d'homologie entre les différents sites centromériques et entre leurs protéines de liaison, cette interaction conduit dans tous les cas

à la formation d'un complexe nucléoprotéique de haut poids moléculaire appelé "complexe de partition ". Le complexe de partition est ensuite reconnu par la protéine motrice et sert de substrat à la réaction de ségrégation. L'architecture et le rôle de ce complexe seront détaillés dans la **Partie III de l'introduction**.

Bien que la fonction générale de la partition soit partagée par les trois types de partition, les mécanismes moléculaires impliqués dans ces systèmes sont très différents et ne sont pas au même stade de compréhension. Dans ce chapitre je détaillerai l'état actuel de nos connaissances sur le fonctionnement et la mécanistique de ces trois types de systèmes.

A. Les systèmes de partition de type I

Les systèmes de partition de type **I** sont les systèmes les plus répandus dans le monde bactérien et les seuls présents sur les chromosomes bactériens. Cependant, les mécanismes moléculaires impliqués n'ont pas encore été élucidés. Les ATPase des systèmes de partition de type **I** font partie d'une famille d'ATPase à motif Walker box A déviant (aussi appelé P-loop) (Koonin, 1993). Ce type est divisé en 2 sous-types selon la taille et l'homologie de séquence de ses protéines, ainsi que sur l'organisation génétique de l'opéron qui les code (Gerdes *et al.*, 2000) (**Figure 12**).

Le sous-type **Ia** est représenté par les deux systèmes de partition les plus étudiés, appelés *parABS* pour le prophage P1 et *sopABC* (*s*tability *o*f *p*lasmid) pour le plasmide F d'*E. coli*. On y retrouve également le système du plasmide RK2/RP4 avec son ATPase, IncC. Les protéines ParA du type **Ia** (ParA_{Ia}) possèdent un double rôle *in vivo*. Elles agissent directement dans le processus de partition (Davis *et al.*, 1996), mais elles jouent aussi un rôle d'autorégulateur capable de réguler la transcription de l'opéron *par* (Mori *et al.*, 1989; T L Davis *et al.*, 1992). Ce rôle de répresseur transcriptionnel est porté par un motif de liaison à l'ADN de type Hélice Tour Hélice (HTH) situé dans la région N-terminale de la protéine. Les protéines ParB_{Ia} contiennent un motif de liaison à l'ADN centromérique, le motif HTH.

Le sous-type **Ib** inclut le système du plasmide pTP228 de *Salmonella newport* avec son ATPase ParF, celui du plasmide pTAR d'*Agrobacterium tumefaciens* et du plasmide pB171 d'*E. coli* avec leur ATPase, ParA. Les protéines ParA et ParB de type **Ib** (ParA_{Ib} et ParB_{Ib}) sont de plus petites tailles. Les protéines ParA_{Ib} ne possèdent pas de motif de liaison à l'ADN, les rendant

incapables de jouer leur rôle de régulateur. Cette fonction est dévolue à ParB_{Ib}. L'opéron étant situé entre les sites centromériques, la fixation de ParB_{Ib} sur ces sites spécifiques lui permet de réaliser l'autorégulation du promoteur (**Figure 12**). Ces protéines ParB_{Ib} ne contiennent pas de motif HTH mais un motif RHH (Ribbon-hélice-hélice), initialement identifié sur la structure de ParB_{Ib} du plasmide TP228 d'*E. coli* (Golovanov *et al.*, 2003). Il est à noter que le plasmide pB171 possède une caractéristique originale. Il contient deux systèmes de partition différents, Par₁ de type **II** et Par₂ de type **I** (Ebersbach and Gerdes, 2001).

Comme énoncé dans le chapitre précédent, la majorité des chromosomes portent également des loci homologues aux loci de partition plasmidiques et tous les systèmes identifiés jusqu'alors sont de type **I**. Cependant ces systèmes apparaissent comme des hybrides entre le type **Ia** et **Ib**. Les protéines ParA sont de type **Ib** alors que les protéines ParB sont très similaires aux protéines ParB_{Ia} avec le motif HTH caractéristique. Les rôles des systèmes de partition chromosomiques ayant été décrits précédemment (**cf. Introduction, Partie I.C**), ils ne seront pas rediscutés ici. Cependant, certaines particularités moléculaires seront tout de même évoquées.

La compréhension des mécanismes moléculaires impliqués dans le phénomène de ségrégation se heurte à la difficulté d'étudier ces deux protéines. En effet, les protéines ParA et ParB sont multifonctionnelles et liées par un jeu de d'inter-régulation, rendant l'interprétation des différents phénotypes observés difficile. Ainsi, l'étude de ces mécanismes moléculaires reste un sujet de recherche qui fait encore débat. Ce débat réside essentiellement au niveau du fonctionnement de la protéine ATPase ParA et comment elle permet la séparation des plasmides. En effet, plusieurs études ont montré que ParA possède un comportement dynamique *in vivo* et *in vitro*. *In vivo*, certaines protéines ParA sont capables d'osciller d'un pôle à l'autre de la cellule. *In vitro*, certaines protéines ParA peuvent former des polymères. Au vu des nombreux résultats recueillis, deux modèles concurrents ont émergé : l'un énonce que l'ATPase ParA, à travers des cycles de polymérisation/ dépolymérisation médiés par son interaction avec le complexe de partition, pourrait provoquer le déplacement des plasmides aux pôles opposés de la cellule, soit en les poussant par extension des filaments, soit en les tirant par dépolymérisation des filaments. Ce modèle est appelé "pulling/pushing". Dans l'autre modèle, ParA se servirait du nucléoïde pour guider le déplacement du plasmide. Dans ce modèle dit de "Réaction-diffusion", la fixation de l'ATPase ParA (en dimères ou petits

oligomères) au nucléoïde serait à l'origine d'un gradient permettant de générer une force pouvant tirer ou pousser les plasmides dans chacune des cellules filles (Hunding *et al.*, 2003).

Les différentes activités de ParA seront tour à tour décrites puis les différents comportements dynamiques de ParA seront plus amplement détaillés afin de comprendre comment ils s'intègrent dans l'un ou l'autre des modèles de ségrégation des plasmides portant les systèmes de partition de type I.

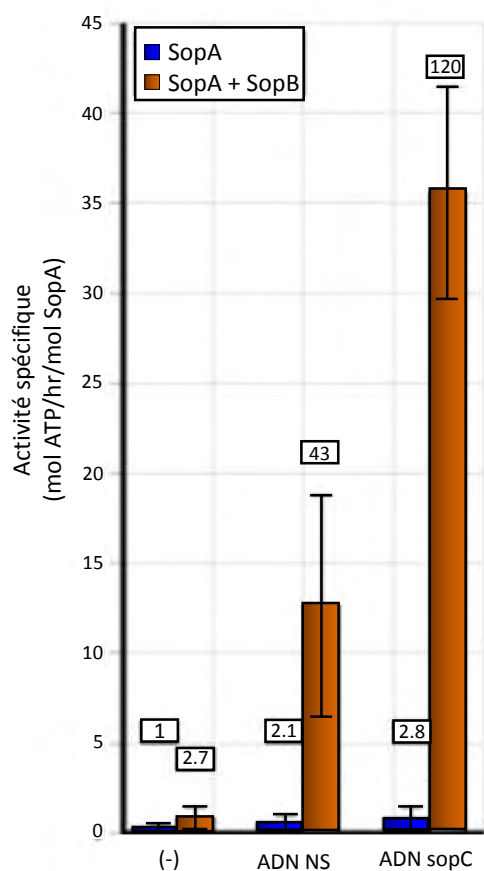
1. L'activité ATPase ParA et autorégulation de l'opéron

Toutes les ATPases de type I possèdent des caractéristiques communes pour leur fonction d'hydrolyse (M A Davis *et al.*, 1992; Watanabe *et al.*, 1992; Leonard *et al.*, 2005; Barillà *et al.*, 2005; Bouet, Ah-Seng, *et al.*, 2007; Pratto *et al.*, 2008). Les alignements de séquences de différentes ParA_{Ia} montrent trois régions similaires impliquées dans la fixation avec l'ATP :

- le motif Walker A est impliqué dans l'interaction avec les phosphates des nucléotides à adénine
- Le motif walker B permet la fixation d'un ion Mg^{2+} , cofacteur de la réaction. Ce motif contient également un résidu aspartate conservé et nécessaire à l'hydrolyse de l'ATP
- Le motif A' possède un résidu aspartate conservé et impliqué dans l'hydrolyse de l'ATP

Des mutations de ces motifs altèrent les propriétés de fixation et d'hydrolyse de l'ATP de ParA (Davis *et al.*, 1996; Fung *et al.*, 2001; Libante *et al.*, 2001; A. G. Vecchiarelli *et al.*, 2013). La présence de la protéine ParB ou d'ADN stimule l'activité d'hydrolyse de ParA respectivement d'un facteur 2 et d'un facteur 3 à 5 environ. La stimulation de cette activité d'hydrolyse est plus importante en présence des deux cofacteurs (facteur 10 à 40). La stimulation de l'ATPase ParF du plasmide TP228 fait exception (ParA_{TP228}). En effet, le niveau de stimulation en présence des deux cofacteurs reste équivalente à celui exercé par ParG seul (ParB_{TP228}) (Barillà *et al.*, 2005). Le rôle du centromère dans la stimulation de l'activité ATPase fait l'objet de résultats contradictoires. Des études réalisées sur les plasmides F et P1 montraient qu'un ADN contenant le site centromérique stimulait l'activité ATPase de la même manière qu'un ADN non spécifique (M A Davis *et al.*, 1992; Watanabe *et al.*, 1992). Or, une étude réalisée sur le système de

a



b

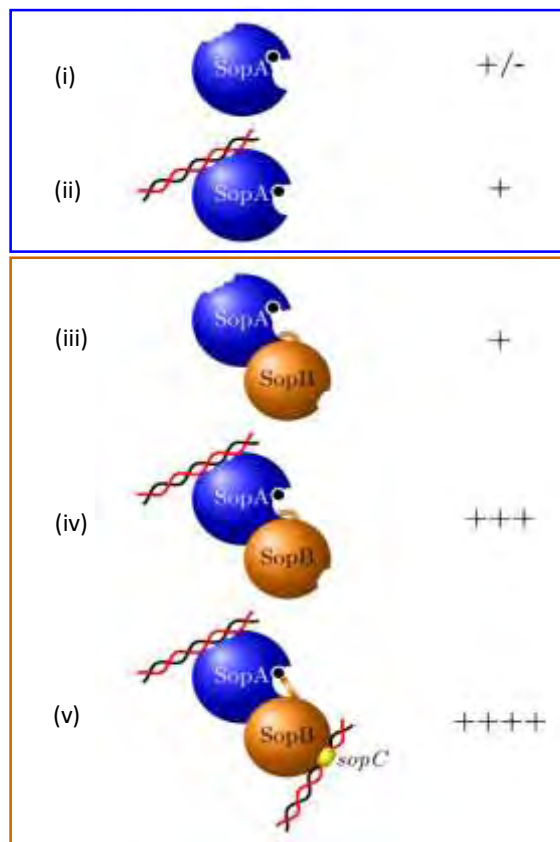


Figure 13 *Stimulation de l'activité ATPase de SopA du plasmide F d'E. coli*

- a) **Caractérisation de l'activité ATPase de SopA.** Les mélanges réactionnels contiennent 2 M de SopA, 2 M de SopB et 60 $\mu\text{g/ml}$ d'ADN non spécifique (NS; pBSKS) ou un plasmide contenant le centromère, *sopC* (pJYB57). Pour chaque valeur mesurée, l'activité ATPase de la même réaction sans SopA, a été soustraite. SopA a été incubée seule (barres bleues) ou avec SopB (barres ocres), en absence d'ADN (-) ou en présence d'ADN NS ou d'ADN contenant *sopC*. Les facteurs de stimulation sont normalisés par rapport à SopA seule (1) et sont présentés à l'intérieur ou au-dessus des barres. Les barres d'erreur ont été obtenues avec 20 expériences indépendantes.
- b) **Modèle de stimulation de l'activité ATPase de SopA par un ADN non spécifique et le complexe de partition.** L'ATPase SopA (sphère bleue) se lie à l'ATP (cercle noir), à l'ADN non spécifique (double hélice noire), et à SopB (sphère ocre). Ces interactions stimulent l'hydrolyse de l'ATP par SopA à des degrés différents, comme indiqué sur la droite. Les protéines SopA et SopB sont sous forme dimérique, mais sont représentées comme des sphères uniques pour plus de simplicité. (i) SopA seule présente une activité ATPase intrinsèque très faible. (ii) en présence d'ADN NS, l'activité ATPase de SopA est modérément activée (facteur 2). (iii) L'interaction avec SopB stimule l'hydrolyse de l'ATP d'un facteur 3 environ. Cette activation nécessite le motif "R-finger" (extension grise qui entre dans la poche catalytique de SopA) aux N-terminale de SopB. (iv) Lorsque les deux ADN NS et SopB sont en contact avec SopA, l'activité ATPase est fortement stimulée (40 fois). Ces deux interactions indépendantes agissent en synergie. SopA liée à l'ADN NS, pourrait adopter une conformation légèrement différente, rendant le motif "R-finger" de SopB plus efficace pour la neutralisation des charges négatives générées lors du transfert de phosphoryle. (v) En présence d'ADN NS et du complexe de partition (SopB-*sopC*), l'activité ATPase de SopA est fortement stimulée (120 fois). Comme dans (iv), les deux interactions indépendantes agissent en synergie et la stimulation supplémentaire d'un facteur 3, pourrait être fournie par un changement de conformation qui se produirait lors de la liaison spécifique à *sopC*. Le motif "R-finger" de SopB pourrait alors adopter une conformation optimale (extension gris droite) dans la poche catalytique de SopA, pour permettre une activité de stimulation maximale. Extrait de (Ah-Seng *et al.*, 2009).

partition de type I du plasmide pSM19035 de *Streptococcus pyogenes* montre qu'en présence du site centromérique *parS* et de ParB_{pSM19035} (ω_2), la stimulation de ParA_{pSM19035} (δ_2), pourrait être deux fois plus importante qu'en présence d'un ADN non spécifique (Pratto *et al.*, 2008). Ceci a été confirmé par une étude sur SopA du plasmide F d'*E. coli*. La stimulation de l'activité ATPase de SopA par SopB et un ADN non spécifique est d'un facteur 40, alors qu'en présence d'un ADN contenant *sopC*, cette stimulation est d'un facteur 120 (facteur trois de différence) (Ah-Seng *et al.*, 2009) (**Figure 13**). Les mécanismes d'activation de l'hydrolyse de l'ATP sont plus compliqués qu'il n'y paraît et les différentes interactions impliquées dans cette stimulation ont pu être précisées ces dernières années.

In vitro, toutes les protéines ParA étudiées sont capables de se lier à l'ADN non spécifique en présence d'ATP (Leonard *et al.*, 2005; Bouet, Ah-Seng, *et al.*, 2007; Hester and Lutkenhaus, 2007; Pratto *et al.*, 2008; Castaing *et al.*, 2008; Hui *et al.*, 2010; Ptacin *et al.*, 2010; Vecchiarelli *et al.*, 2010). Un mutant de SopA (SopA-K340A), incapable de reconnaître l'ADN, ne stimule plus efficacement l'hydrolyse de l'ATP (Castaing *et al.*, 2008), suggérant un rôle de cette interaction dans l'hydrolyse de l'ATP. Par ailleurs, cette stimulation peut également se faire via ParB. Il a été montré que la stimulation de plusieurs NTPases se faisait par l'intermédiaire d'un motif arginine "finger" (Scheffzek, 1997; Rittinger *et al.*, 1997; Scheffzek *et al.*, 1998). L'arginine stabilise l'état de transition de la réaction d'hydrolyse et permet de faire le changement de conformation de la protéine après hydrolyse du nucléotide (Scheffzek, 1997; Rittinger *et al.*, 1997; Nadanaciva *et al.*, 1999). Une arginine "finger" capable de stimuler l'activité ATPase des protéines ParF et SopA, a pu être identifiée respectivement sur ParG du plasmide pTP228 et SopB du plasmide F (Barillà *et al.*, 2007; Ah-Seng *et al.*, 2009). Pour SopB cette arginine est en position 36 et pour ParA_{TP228} en position 19. Ah-Seng et ses collaborateurs ont montré que pour obtenir une activation maximale d'un facteur 120, la stimulation de l'activité ATPase devait se faire via deux interfaces (**Figure 13**). SopA interagit d'une part avec l'ADN non spécifique puis avec SopB préalablement lié à de l'ADN contenant *sopC*. L'identification du domaine d'interaction de SopB impliqué dans la liaison avec SopA et le rôle de cette interaction dans la régulation des activités de SopA seront détaillés dans ce manuscrit (**cf. Résultat, Chapitre IV**).

Le rôle exact de la fixation et de l'hydrolyse de l'ATP dans le processus général de partition n'est pas encore compris cependant cette activité est importante pour la partition puisqu'un mutant inhibant l'hydrolyse de l'ATP, entraîne un défaut de partition (Davis *et al.*,

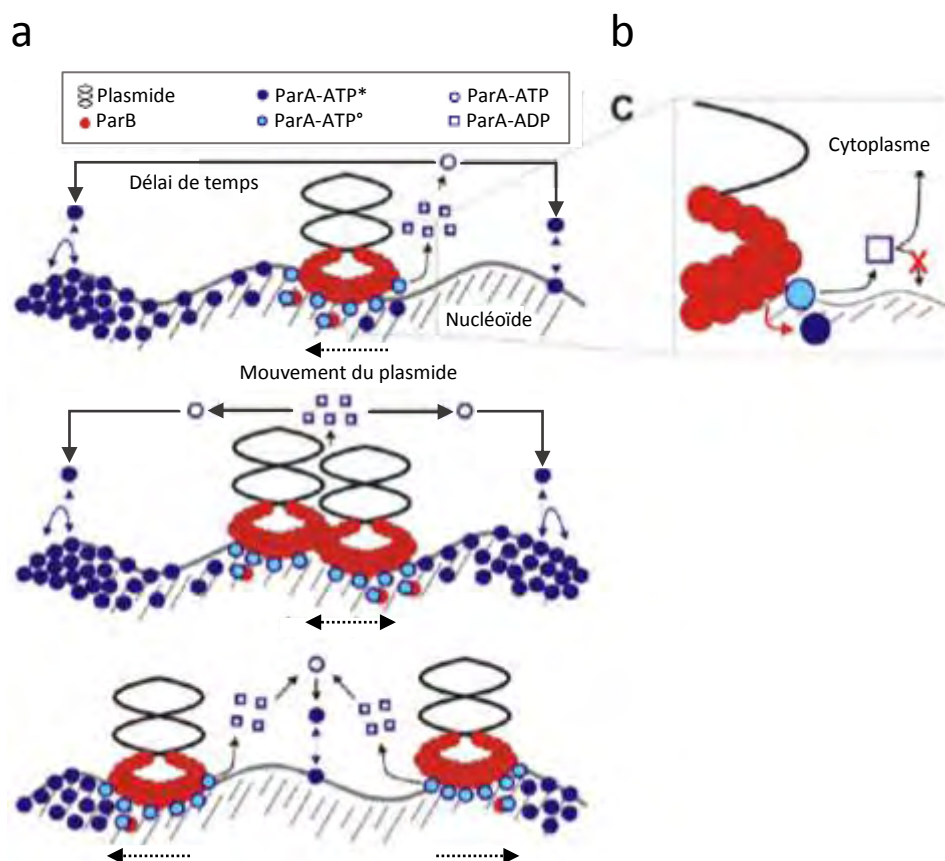


Figure 14 | Le modèle de "réaction-diffusion"

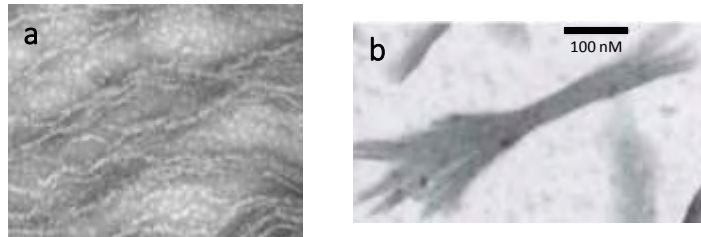
- a) **Transport du plasmide P1 par l'intermédiaire du mouvement de *parA* sur le nucléoïde.** ParA se lie à l'ATP et pourrait adopter une conformation avec laquelle ParA serait capable de se lier au nucléoïde, ParA-ATP*. Ainsi, ParA-ATP* pourrait se lier au nucléoïde et les dimères de ParB se fixeraient sur le plasmide pour former le complexe de partition (surface hachurée). L'interaction entre ParA et ParB lié au nucléoïde, pourrait ancrer le plasmide sur le nucléoïde. L'activité ATPase de ParA serait stimulée par ParB. Cette stimulation pourrait convertir ParA-ATP* en une forme qui hydrolyse l'ATP, ParA-ATP°, créant ainsi une zone de déplétion sur le nucléoïde tout autour du plasmide. ParA-ADP se dissocierait du nucléoïde. Puis, ParA pourrait échanger son ADP pour l'ATP (ParA-ATP) avec un délai de temps pour retrouver la conformation de ParA-ATP*. Ce délai permettrait à ParA de diffuser de manière aléatoire avant de se réassocier au nucléoïde. La redistribution continue de ParA sur le nucléoïde, pourrait permettre de pousser le plasmide. Une fois la réplication effectuée, les plasmides se sépareraient de manière bidirectionnelle comme s'ils suivaient les concentrations élevées de ParA lié au nucléoïde.
- b) **Modèle du déplacement du plasmide.** L'interaction entre ParA-ParB produit du ParA-ADP, qui ne pourrait plus rester lié au complexe de partition ou au nucléoïde. Ainsi, ParA-ADP serait libérée dans le cytoplasme. ParB en revanche resterait pour se lier immédiatement à un dimère de ParA lié au nucléoïde. Extrait de (Vecchiarelli *et al.*, 2012).

1996; Fung *et al.*, 2001; Libante *et al.*, 2001; Barilla *et al.*, 2007). Des modèles de partition impliquant l'hydrolyse de l'ATP, ont été développés. L'un d'eux décrit un mécanisme de "réaction-diffusion" et prend en compte le cycle de l'ATP, avec la fixation de l'ATP et son hydrolyse.

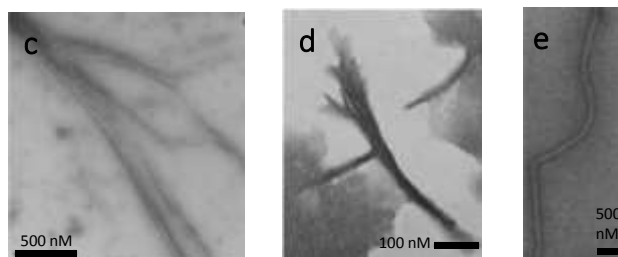
Ce modèle de "réaction-diffusion" a été précisé à l'échelle moléculaire par une étude récente sur le plasmide P1 (**Figure 14**). Une première étape clé de ce mécanisme, est le changement de conformation dépendante de l'ATP, nécessaire pour convertir ParA en sa forme liant l'ADN non spécifique (Vecchiarelli *et al.* 2010). Cette étape est très lente, ce qui, *in vivo*, créerait un délai entre le moment où ParA se détache du nucléoïde après hydrolyse de l'ATP, et le moment où ParA acquiert de nouveau la forme qui se lie au nucléoïde. Ce délai, pourrait permettre à ParA de diffuser à travers la cellule et de se réassocier aussi bien à proximité du lieu où elle a été libérée qu'à n'importe quel autre endroit sur le nucléoïde et ceci avec la même probabilité. La stimulation de l'activité ATPase de ParA est maximale au niveau du complexe ParB-plasmide, entraînant le décrochage de ParA de l'ADN. Ce décrochage, couplé au délai de temps pour réaliser la réassociation de ParA à l'ADN, conduit à une redistribution continue de ParA le long du nucléoïde. Cette redistribution crée un gradient de ParA sur le nucléoïde qui pourrait être à l'origine du mouvement des plasmides. Dans ce modèle de "réaction-diffusion", les plasmides migrent le long du nucléoïde vers une concentration locale élevée de ParA. Ainsi, entraînés par cette vague de ParA, les plasmides pourraient "surfer" sur le nucléoïde. La clé de ce modèle réside dans le délai entre les différents changements de conformation.

Le cycle de l'ATP est donc un déterminant majeur dans le comportement dynamique des ATPases de type Walker box impliquées dans le mécanisme de partition. Pour certaines protéines ParA, appartenant au sous-type Ia, ce cycle de l'ATP est également important pour la fonction de régulation de l'opéron *par* en plus de son activité dans le mécanisme de partition. Par (P1) et SopA (F), se fixent sur les sites opérateurs de leur région promotrice *parOP* et *sopOP* par l'intermédiaire de leur motif HTH de liaison à l'ADN (Mori *et al.*, 1989; T L Davis *et al.*, 1992; Hayes *et al.*, 1994; Ravin *et al.*, 2003). Le passage entre la forme impliquée dans la partition et celle dans la répression est dépendante du nucléotide ATP ou ADP fixé. En effet, la forme ParA-ATP ou SopA-ATP impliquée dans la partition, n'est pas la forme répresseur. Pour P1, la forme ParA-ADP est la forme répresseur (Davey and Funnell, 1994; Bouet and Funnell, 1999; Fung *et al.*, 2001) et pour F, c'est la forme SopA-ADP ou SopA-libre (Mori *et al.*, 1989; Bouet, Ah-Seng,

ATPase de type Ia



ATPase de type Ib



ATPase de type Ic

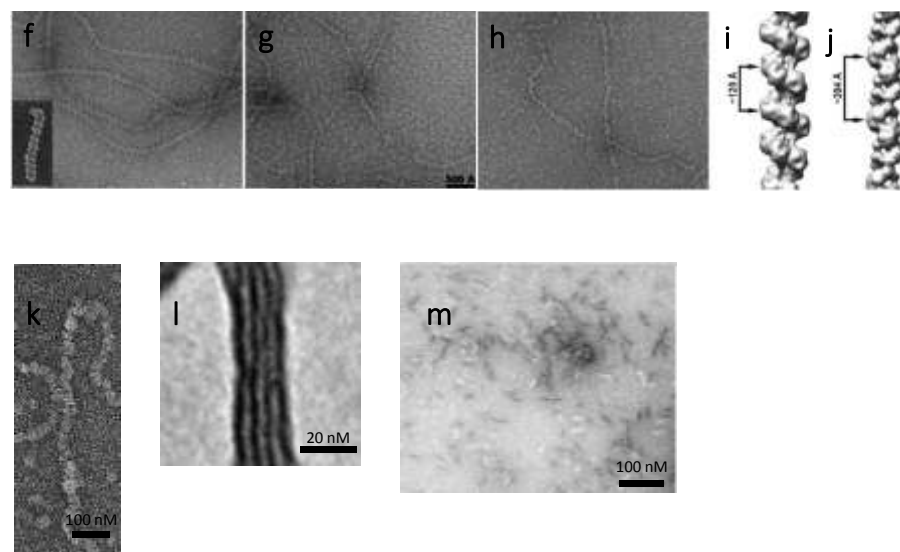


Figure 15 *∫* Filaments des ParA plasmidiques et chromosomiques de type I, visualisés *in vitro* en ME après coloration négative

Les images sont classées selon les 3 sous-types des systèmes de partition de type I et sont indiqués en gris. Les protéines ParA des plasmides P1 (a), F (b), pB171 (c), pTP228 (d), forment des filaments en présence d'ATP *in vitro*. Les filaments de $\delta 2$ (ParA) de pSM19035 se forment en présence d'ATP et d'ADN centromérique (e). Les images de f à i représentent des filaments de ParA₂ du chromosome 2 de *V. cholerae* sur de l'ADN double brin, en présence d'ATP (f), d'ADP (g) ou d'aucun nucléotide. (h) et (i) sont des reconstructions tridimensionnelles des filaments de ParA₂ formés sur l'ADN en présence d'ATP (h) ou d'ADP (i). (k) Les filaments de Soj de *T. thermophilus* ont une forme régulière et hélicoïdale. Ils se forment en absence de nucléotides et sur un ADN linéarisé. (l) Les filaments de ParA de *C. crescentus* forme des structures fibreuses. (m) L'ATPase SegA de *S. solfataricus*, une archée, peut également former des polymères en présence d'ATP.

Extrait de (Leonard *et al.*, 2005; Barilla *et al.*, 2005; Ebersbach *et al.*, 2006; Bouet, Ah-Seng, *et al.*, 2007; Pratto *et al.*, 2008; Dunham *et al.*, 2009; Hui *et al.*, 2010; Ptacin *et al.*, 2010; Kalliomaa-Sanford *et al.*, 2012).

et al., 2007). *In vivo*, ParB/SopB augmente fortement la répression transcriptionnelle de ParA/SopA (Friedman and Austin, 1988; Hirano *et al.*, 1998) et *in vitro*, ParB/SopB stimule la fixation de l'ATPase à leur région opératrice (Mori *et al.*, 1989; Davey, 1997). L'ensemble des résultats obtenus et le fait que ParB/SopB stimule l'hydrolyse de l'ATP, suggère que l'une des fonctions de ParB/SopB est de convertir ParA en sa forme répresseur. Les centromères *parS/sopC* sont aussi nécessaires dans l'activité de répression du promoteur (Yates *et al.*, 1999; Hao and Yarmolinsky, 2002). Leur effet n'est visible qu'en présence de la protéine ParB/SopB. Ces données suggèrent que le complexe de partition (ParB-centromère), stimule l'activité ATPase et permet de convertir plus efficacement ParA/SopA dans sa forme répresseur.

Une étude sur ParA a montré que l'autorégulation du promoteur, n'avait pas d'effet sur la partition du plasmide (Davis *et al.*, 1996). Ainsi, l'autorégulation de la transcription de l'opéron *par* ne pourrait jouer qu'un rôle de soutien dans la ségrégation plasmidique en veillant à ce que les niveaux des protéines codées restent dans des limites fonctionnelles. Les rôles exacts de l'activité ATPase et de l'hydrolyse en tant que tels dans le mécanisme de partition sont encore sujet à débat.

2. Le comportement dynamique de ParA

a) Les polymères de ParA

Les ATPases de type I sont capables de former des polymères *in vitro*. Ces polymères ont pu être observés, aussi bien pour des ATPase plasmidiques que pour des ATPases chromosomiques (**Figure 15**). Des structures apparentées ont également pu être décrites *in vivo*. Par exemple, ParA de *C. crescentus*, visualisée en microscopie à fluorescence, forme des structures linéaires *in vivo* (Ptacin *et al.*, 2010). Une étude sur le plasmide F a permis d'attribuer un rôle des polymères de SopA dans la ségrégation du plasmide (Lim *et al.*, 2005). Dans cette étude, les auteurs ont pu suivre la cinétique d'assemblage de chaque filament de SopA en microscopie à fluorescence en "Time-lapse". L'allongement des polymères de SopA se fait à un taux moyen de 0,18 $\mu\text{m}/\text{min}$. Cette vitesse est comparable à la vitesse de séparation des plasmides ou des origines bactériennes *in vivo*. Par exemple, le taux moyen de séparation pour les plasmides F, RK2, et P1 est de 0,43, 0,3 et 0,41 $\mu\text{m}/\text{min}$, respectivement, et pour la séparation des origines des chromosomes de *C. crescentus*, *B. subtilis* et *E. coli*, la vitesse de séparation est en moyenne de 0,27, 0,17, et 0,07 $\mu\text{m}/\text{min}$, respectivement (Webb *et al.*, 1998;

| Type de système | ParA | [parA] (μM) | Nucléotide requis | Effet des cofacteurs sur la polymérisation | Techniques | Références |
|-----------------|-----------------------------|-------------|--------------------------------|--|-------------------------------------|---|
| Ia | ParA plasmide P1 | 0,5-40 | ATP ATPyS | Stimulation par ParB et l'ADN | DLS Sédimentation ME | (Havey <i>et al.</i> , 2012) (Dunham <i>et al.</i> , 2009) |
| | SopA plasmide F | 1-20 | ATP | Stimulation par SopB et inhibé par l'ADN | DLS Sédimentation ME, M-Fluo* | (Lim <i>et al.</i> , 2005) (Bouet <i>et al.</i> , 2007) |
| Ib | ParA pB171 | 15-20 | ATP ATPyS* | Stimulation par ParB (ADN non testé) | DLS Sédimentation ME | (Ebersbach <i>et al.</i> , 2006) (Machón <i>et al.</i> , 2007) |
| | ParF plasmide TP228 | 2-10 | ATP ATPyS | Stimulation par ParB (ADN non testé) | DLS Sédimentation ME | (Barillà <i>et al.</i> , 2005) (Barillà <i>et al.</i> , 2007) (Machón <i>et al.</i> , 2007) |
| | δ (ParA) pSM19035 | 1 | ATP ATPyS | Stimulation par ω (ParB) et l'ADN | DLS | (Pratto <i>et al.</i> , 2008) |
| | ParA pTAR & pVT745 | 2-10 | ATP ATPyS | Stimulation par ParB (ADN non testé) | DLS Sédimentation | (Machón <i>et al.</i> , 2007) |
| Chromosomique | ParA2 <i>V. cholerae</i> | 1-10 | ATP ADP aucun nucléotide | ADN requis (ParB et ADN non testés) | DLS ME | (Hui <i>et al.</i> , 2010) |
| | Soj <i>T. thermophilus</i> | 25 | ATP | ADN requis (ParB non testé) | DLS ME | (Leonard <i>et al.</i> , 2005) |
| | ParA <i>C. crescentus</i> | 1-40 | ATP | Pas de données | DLS ME | (Ptacin <i>et al.</i> , 2010) |
| | SegA <i>S. solfataricus</i> | 5 | ATP ATPyS ADP | Stimulation par SegB (ADN non testé) | DLS Sédimentation ME | (Kalliomaa-Sanford <i>et al.</i> , 2012) |

*ATPyS induit la polymérisation de ParA de pB171 dans (Manchon *et al.*, 2007), mais pas dans (Ebersbach *et al.*, 2006)

*microscopie à fluorescence

Tableau 3 | Variations autour des conditions de polymérisation entre les différents systèmes de type I

Certaines protéines ParA plasmidiques et chromosomiques sont capables de polymériser en faisceaux de filaments, *in vitro*. Ce tableau regroupe les variations de conditions de polymérisation des protéines entre les systèmes par. Certaines protéines polymérisent avec de l'ATPyS, ADP ou sans aucun nucléotide. Les effets de la protéine ParB et de l'ADN n'ont pas été testés dans tous les cas. Pour ceux qui l'ont été, la présence de ParB et/ou de l'ADN sont requis ou stimulent la polymérisation de ParA, à l'exception de SopA pour lequel l'ADN inhibe la polymérisation.

Adaptée de (Vecchiarelli *et al.*, 2012).

Pogliano *et al.*, 2001; Viollier *et al.*, 2004; Gordon *et al.*, 2004; Gitai *et al.*, 2005). Ces résultats ont permis de proposer un modèle dans lequel la polymérisation permettrait de pousser les plasmides à l'autre extrémité de la cellule. Cependant, la forme des filaments ainsi que les conditions de polymérisation varient d'un système *par* à l'autre, rendant leur rôle exact dans la partition difficile à cerner (**Figure 15, Tableau 3**).

(1) Rôle de l'ATP dans la polymérisation de ParA

L'ATPase ParB_{TP228} (ParF) forme des polymères en présence d'ATP et d'ATPγS, suggérant que l'hydrolyse de l'ATP est nécessaire pour la formation des polymères (Barillà *et al.*, 2005; Havey *et al.*, 2012). Au contraire, SopA de F forme des filaments en présence d'ATP mais pas en présence d'ATPγS (Bouet, Ah-Seng, *et al.*, 2007). Deux explications sont possibles, soit l'hydrolyse de l'ATP est nécessaire pour la formation de polymères, soit la conformation adoptée par SopA en présence d'ATPγS ne permet pas l'assemblage des polymères. Une étude a permis de montrer que la polymérisation de SopA requière une conformation précise avec l'ATP qui n'est pas possible avec l'ATPγS (Bouet, Ah-Seng, *et al.*, 2007). Dans le cas des ATPases chromosomiques, seule l'ATP est essentielle à la polymérisation de ParA (Soj) de *T. thermophilus* et *C. crescentus* (Leonard *et al.*, 2005; Ptacin *et al.*, 2010). Pour ParA2 de *V. cholerae* (ParA2_{V.ch}), la polymérisation se fait en présence de ParA2 seule, d'ADP ou d'ATP, alors que pour *S. sulfataricus* elle se fait en présence d'ATP, ADP ou d'ATPγS.

A l'exception de ParA2_{V.ch}, toutes les protéines ParA ont besoin d'un nucléotide à adénine pour pouvoir polymériser. Toutes les protéines ParA testées sont capables de polymériser en présence d'ATPγS à l'exception de SopA. Ainsi, l'hydrolyse de l'ATP en tant que telle pourrait ne pas être nécessaire au mécanisme de polymérisation alors que la conformation adoptée lorsque ParA est liée à un nucléotide à adénine pourrait l'être.

(2) Rôle de ParB et de l'ADN dans la polymérisation de ParA

Dans tous les cas où cela a été testé, la présence de la protéine ParB stimule la polymérisation de l'ATPase. De même, l'ADN permet de stimuler cette polymérisation et parfois sa présence est absolument requise (Leonard *et al.*, 2005; Hui *et al.*, 2010) (**Figure 15**). SopA fait exception. En effet, l'ADN inhibe sa polymérisation (Bouet, Ah-Seng, *et al.*, 2007). *In vivo*, étant donné la présence importante de l'ADN dans la cellule, on pourrait penser que la formation de ces polymères soit inhibée. Or, des études de microscopie montrent que SopA

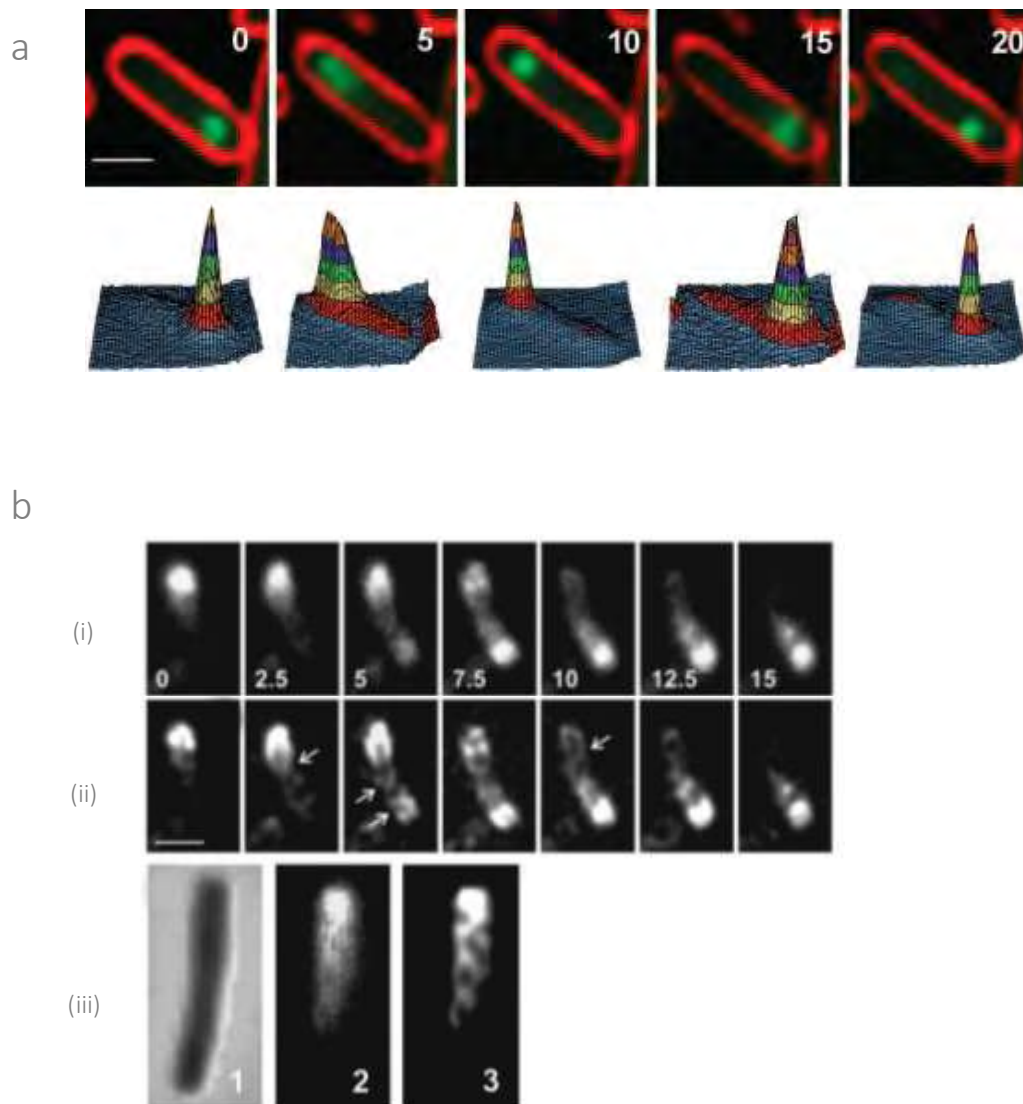


Figure 16 *Comportement dynamique des ParA plasmidiques et chromosomiques, in vivo.*

- a) *Oscillation de SopA-GFP du plasmide F, visualisée en microscopie time-lapse.* (haut) Les membranes cellulaires sont colorées avec FM 4-64 (rouge). Le temps en minutes où les acquisitions ont été effectuées, est indiqué en blanc dans le coin supérieur droit de chaque panneau. La barre correspond à 1 μ m. (bas) Ces diagrammes représentent l'intensité de fluorescence de SopA-GFP en fonction du temps et de sa localisation dans la cellule. Les couleurs indiquent l'intensité des pixels sur une échelle linéaire et en unités arbitraires : bleu (0-200), rouge (200-400), jaune (400-600), vert (600-800), violet (800-1000) et orange (1000-1200).
- b) *Oscillation et structure hélicoïdale de ParA de pB171.* (i) Suivi de ParA-GFP de pB171 pendant 15 min avec une prise d'image toute les 30s (ii) Les images A sont soumises à un programme de déconvolution 2D. Les flèches montrent les croisements et des structures hélicoïdales fluorescentes. (iii) Prise d'une série d'images en fonction de l'axe Z de la cellule. "1" représente une cellule en contraste de phase soulignant les contours de la cellule, "2" représente la superposition de toutes les images prises selon l'axe Z et en "3", l'image "2" est soumise au programme de déconvolution 3D révélant une structure hélicoïdale que dessine Par-GFP lors de son déplacement. Extrait de (Ebersbach and Gerdes, 2004).

forme des structures fibreuses *in vivo*, suggérant ainsi que la formation de ces polymères est possible (Lim *et al.*, 2005; Hatano *et al.*, 2007). *In vitro*, la présence d'une protéine de liaison à l'ADN comme SopB permet de rétablir la polymérisation de SopA en présence d'ADN (Bouet, Ah-Seng, *et al.*, 2007). Nous verrons par la suite comment SopB par l'intermédiaire de tout un jeu d'interactions pourrait masquer l'ADN et créer les conditions nécessaires à la polymérisation de SopA.

Bien que la polymérisation des ATPases de type I en filaments a pu être observée à la fois *in vivo* et *in vitro*, leur capacité à réaliser directement la ségrégation des plasmides, n'a pas encore été démontrée. De plus, l'hétérogénéité au niveau des conditions de polymérisation des différents ParA, rend très difficile la compréhension de la régulation de la formation de ces polymères ainsi que le rôle que l'on peut leur accorder.

b) L'oscillation de ParA

Certaines versions fluorescentes des protéines ParA (ParA de P1, SopA de F, ParA de pB171, de pSM19035 et les ParA chromosomiques de *V. cholerae* (ParAI) et *B. subtilis* (Soj)), ont un comportement dynamique qui colocalise avec le nucléoïde (Marston and Errington, 1999; Quisel *et al.*, 1999; Lim *et al.*, 2005; Ebersbach *et al.*, 2006; Fogel and Waldor, 2006; Hatano *et al.*, 2007; Pratto *et al.*, 2008; Ringgaard *et al.*, 2009; Hatano and Niki, 2010) (**Figure 16**). Cependant, en absence de ParB, ParA est diffus sur le nucléoïde et le comportement dynamique est perdu.

L'une des premières études *in vivo* sur le comportement oscillatoire des protéines ParA plasmidiques ont été faites sur le système Par₂ de type I du plasmide pB171 d'*E. coli* (Ebersbach and Gerdes, 2001; Ebersbach and Gerdes, 2004). Les auteurs ont montré que ParA_I de pB171, fusionnée à la GFP, oscille plusieurs fois par cycle cellulaire, sur une région restreinte au nucléoïde. De manière similaire, SopA du plasmide F a aussi un comportement dynamique *in vivo*, se traduisant aussi par une oscillation sur le nucléoïde (Lim *et al.*, 2005; Adachi *et al.*, 2006; Hatano *et al.*, 2007) (**Figure 16a**). Au cours du temps, l'amas de SopA voyage d'un pôle à l'autre de la cellule. Dans les deux cas, l'utilisation d'un programme de déconvolution a permis de visualiser que cette oscillation décrit un mouvement en forme de spirale autour du nucléoïde (Ebersbach and Gerdes, 2004; Adachi *et al.*, 2006). L'oscillation de ParA et SopA est dépendante

de ParB/SopB et *parS/sopC* respectivement (Ebersbach and Gerdes, 2001; Adachi *et al.*, 2006), suggérant que le complexe de partition régule les propriétés dynamiques de ces protéines.

Cette régulation semble venir de la capacité du complexe de partition à réguler l'activité ATPase des protéines ParA. En effet, un mutant de ParB (ParB_{26A}) qui ne stimule plus l'activité ATPase de ParA de *par*₂ de pB171, ne possède plus de comportement dynamique (Ringgaard *et al.*, 2009). De même, un mutant de SopA directement déficient pour son activité ATPase (hydrolyse (K₁₂₀R) ou fixation de l'ATP (K₁₂₀Q)), n'oscille plus et se localise de manière diffuse dans la cellule (Hatano *et al.*, 2007). Une corrélation entre l'oscillation des ParA, l'activité ATPase et la stabilité des plasmides a pu être montrée grâce à une mutation d'un résidu conservé de liaison à l'ATP, dans le motif Walker box de ParA du système *par*₂ de pB171, suggérant un rôle de l'oscillation de ParA, dans la ségrégation des plasmides (Ebersbach and Gerdes, 2001).

SopB forme des foci qui colocalisent avec le plasmide F. Bien qu'une petite fraction des protéines SopB a été observée sous forme de spirale (Adachi *et al.*, 2006), cette observation n'a été confirmée sur aucun autre système de partition (Lim *et al.*, 2005) ni même sur F (Sanchez *et al.*, 2013). Le suivi de plasmide F et de SopA en microscopie et en time lapse montre que le plasmide semble suivre l'amas de SopA. L'oscillation de SopA pourrait recruter SopB lié au plasmide et participerait à sa localisation aux positions $\frac{1}{4}$ et $\frac{3}{4}$ de la cellule (Hatano *et al.*, 2007).

L'oscillation des protéines ParA est compatible avec le modèle de "réaction-diffusion". Dans ce modèle, la diffusion des SopA matérialisée par cette oscillation permettrait de recruter et disposer SopB lié au plasmide de manière ordonnée. Cependant, l'oscillation est aussi compatible avec le modèle des filaments dit "pulling", dans lequel ParA-ATP polymérise en formant des filaments hélicoïdaux. Quand une extrémité du filament rencontre les protéines ParB liées au plasmide, la réaction s'inverse et le filament se dépolymérise, tirant le plasmide sur le nucléoïde vers le pôle opposé (Ringgaard *et al.*, 2009; Ptacin *et al.*, 2010). Cette observation faite sur le plasmide pB171_{par2} et F, peut aussi être étendue aux systèmes de partition chromosomiques puisque des résultats similaires ont été obtenus avec la protéine ParA du chromosome I de *V. cholerae* (Fogel and Waldor, 2006).

Ce mécanisme d'oscillation est aussi utilisé pour positionner de larges complexes protéiques (carboxysomes et large clusters de chimiotaxie) (Szardenings *et al.*, 2011). La capacité de toutes ces protéines à osciller et former des structures hélicoïdales dans les cellules a récemment été remise en question par l'étude d'une protéine responsable de la morphologie en bâtonnet de quelques espèces bactériennes, MreB (pour revue (White and Gober, 2012)). Cette protéine homologue des actines, a été la première protéine décrite comme étant capable de former des longs filaments hélicoïdaux traversant les bactéries en bâtonnet (Jones *et al.*, 2001) et serait un composant majeur du cytosquelette bactérien. Récemment, plusieurs équipes ont démontré comment la protéine MreB, qui apparaît sous forme de foci par la technique de microscopie TIRFM (Total Internal Reflection Fluorescence Microscopy), peuvent être interprétés à tort comme de longs filaments en microscopie à épi-fluorescence classique. En effet, les fusions fonctionnelles de MreB-GFP forment des foci discrets et dynamiques sur toute la circonférence de la cellule et non des structures en spirale (Dominguez-Escobar *et al.*, 2011; Garner *et al.*, 2011). Ces études posent la question de l'intégrité de certaines structures spiralées décrites, dont les protéines ParA font parties. Ainsi, l'ATPase de type **Ib**, Soj, codée par le chromosome de *B. subtilis* avait été initialement décrite comme étant capable d'osciller *in vivo* (Marston and Errington, 1999; Quisel *et al.*, 1999). Ce mouvement dynamique était dépendant de la présence de l'homologue de ParB, SpoOJ. En absence de celui-ci, Soj se localise de manière non spécifique sur le nucléoïde comme cela a été vu pour ses homologues. Dans une étude ultérieure réalisée dans des conditions proches de l'expression physiologique, aucun comportement oscillatoire de la protéine Soj n'a pu être observé et Soj forme des foci sur tout le long de la cellule (Murray and Errington, 2008). Par ailleurs, il a récemment été montré en cryo-électron-tomographie et en microscopie à fluorescence que le nucléoïde de l'ensemble des bactéries Gram-négatives et Gram-positives pouvait adopter une ultrastructure particulière organisée en spirale (Berlatzky *et al.*, 2008; Butan *et al.*, 2010). Ainsi, les structures en spirales adoptées par MreB ou ParA, pourraient n'être que le résultat de la diffusion d'un amas concentré de protéines entre la membrane interne de la cellule et le nucléoïde.

Au vu de ces dernières données, l'implication de l'oscillation en tant que structure en spirale dans le mécanisme de partition est discutable ainsi que le modèle qui en résulte. Cependant, de plus en plus d'évidences apparaissent sur le rôle du nucléoïde bactérien dans la réaction de partition. Il pourrait servir de matrice à la diffusion des protéines ParA qui

entraînerait le complexe ParB-plasmide dans leur déplacement. En effet, les mouvements des amas de ParA sont restreints au nucléoïde et les mouvements des plasmides sont eux aussi localisés à cet endroit (Erdmann *et al.*, 1999; Ebersbach and Gerdes, 2004; Derman *et al.*, 2008; Ringgaard *et al.*, 2009). De plus, ParA possède une activité de liaison à l'ADN non spécifique et toute mutation affectant l'activité de liaison à l'ADN conduit à une perte du plasmide (Hester and Lutkenhaus, 2007; Castaing *et al.*, 2008). Pour mieux comprendre un rôle éventuel du nucléoïde dans la partition, deux équipes ont récemment reconstitué la réaction de partition du plasmide F et P1, dans un système acellulaire. Pour recréer un nucléoïde artificiel, ils ont reconstitués un tapis d'ADN non spécifique sur la surface d'une lame de microscopie (Hwang *et al.*, 2013). Par un système de microfluidique, les protéines ParA et ParB fluorescentes sont injectées sur un tapis d'ADN puis visualisées en microscopie TIRFM. Pour les deux systèmes, plasmides F et P1, il apparaît que l'ADN situé à proximité des clusters de complexes de partition, est déplété en ATPase. Ces zones de déplétion sont une composante essentielle du modèle de "réaction-diffusion" (Vecchiarelli *et al.*, 2010). Ils ont observé des comportements dynamiques et des mouvements latéraux pour de nombreux clusters de complexes de partition à la fois pour P1 et F (Hwang *et al.*, 2013; A. G. Vecchiarelli *et al.*, 2013) ce qui appuie leur modèle. Cependant, après la formation de la zone de déplétion de l'ATPase, la plupart des complexes de partition se dissocient du tapis d'ADN sans aucun mouvement latéral. L'explication avancée par les auteurs est que leur système expérimental de microfluidique ne permet pas de recréer les conditions exactes de confinement présentes dans le cytoplasme. Or dans la cellule, l'espace restreint entre le nucléoïde et la membrane cytoplasmique pourrait contraindre les complexes de partition à rester à proximité du nucléoïde (Vecchiarelli *et al.*, 2014). Ainsi après dissociation, les complexes se réassocient immédiatement à proximité du lieu de décrochage et réalisent des mouvements latéraux comme décrit dans le modèle de "réaction-diffusion".

Dans les deux expériences précédentes, il est à noter qu'aucun polymère à proprement dit, n'a pu être détecté pour ParA de P1 ou SopA de F. Aussi, ParA et SopA pourraient utiliser le même principe de "réaction-diffusion" pour accomplir le même résultat biologique, à savoir l'héritage fidèle de l'information génomique. Toutefois des différences entre les systèmes ParABS de P1 et SopABC de F existent et peuvent inclure des différences de concentration des protéines de partition *in vivo*, des différences quantitatives au niveau de différents paramètres cinétiques, ou d'éventuelles interférences entre les systèmes de partition et leur

environnement *in vivo*. Ces subtiles différences dans leurs mécanismes biochimiques respectifs, expliquent probablement les variations du comportement des ATPase observées *in vivo*. En effet, la protéine ParA de P1 forme des foci intenses et clignotant, colocalisés avec le plasmide P1 (Hatano and Niki, 2010). ParA migre progressivement du milieu de la cellule à la position $\frac{1}{4}$ et $\frac{3}{4}$ de la cellule alors que SopA du plasmide F oscille sur le nucléoïde *in vivo*. La localisation et le positionnement du plasmide P1 semblent se faire entre les phases mobiles et immobiles de la protéine ParA (Hatano and Niki, 2010). Ainsi, la distribution d'un gradient de ParA détermine la destination des plasmides. Pour F, c'est le comportement oscillatoire de SopA en tant que gradient protéique mobile, qui pourrait permettre cette étape de positionnement (Hatano *et al.*, 2007). Le modèle n'exclut pas non plus un rôle de la polymérisation des ParA. Par exemple, les amas de ParA liés à l'ADN, observés dans les expériences de TIRFM pourraient être des dimères ou de courts oligomères.

La principale différence entre le modèle de "pulling" principalement représenté par les plasmides de type **Ib**, et de "réaction-diffusion" par ceux de type **Ia**, se résume à comment ParA agit pour diriger la ségrégation des plasmides : dans le premier modèle, ParA forme des filaments qui à eux seul porte le mécanisme de ségrégation (Ringgaard *et al.*, 2009), et dans le second, ParA forme un gradient dynamique de dimères ou de petits oligomères (Vecchiarelli *et al.*, 2010). Les modèles ont de grandes similitudes, et il semble probable que chacun de ces mécanismes soit dans le fond, essentiellement le même, même s'ils diffèrent pour quelques détails (Howard and Gerdes, 2010).

B. Les systèmes de partition de type II

Le système de partition de type **II** est le système de partition le mieux compris d'un point de vue moléculaire bien qu'il soit très peu répandu. Le locus *parRMC* a été découvert il y a 25 ans sur le plasmide R1 d'*E. coli* (Gerdes *et al.*, 1985). Des homologues de ce système de partition de type **II** sont seulement présents sur quelques plasmides des bactéries Gram-positives ou Gram négatives et aucun locus *parRMC* n'a pu être identifié sur les chromosomes bactériens (Ebersbach and Gerdes, 2001; Becker *et al.*, 2006; Schumacher, 2007). Ce locus code pour une protéine homologue aux actines/hsp70 (heat-shock protein 70) eucaryotes, une protéine adaptatrice ParR et une séquence centromérique *parC*. Comme pour le système de type **Ib**, le gène *parR* est situé en aval du gène *parM* et le site centromérique en amont du locus

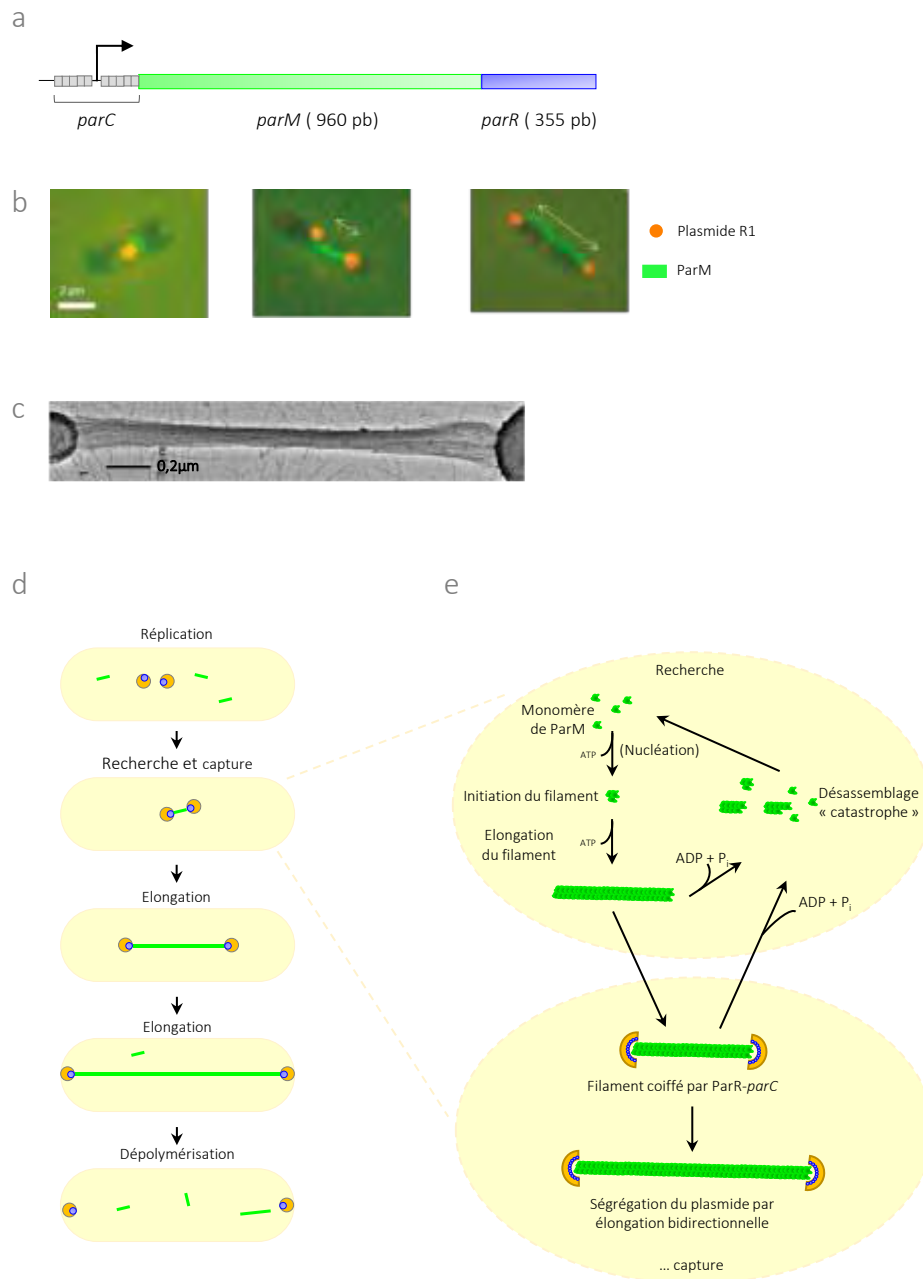


Figure 17 | Modèle de "pushing" du mécanisme de la ségrégation du plasmide R1

- Organisation génétique.** Le centromère est composé de deux blocs de cinq courtes séquences répétées. Les deux blocs sont séparés par 39 pb contenant le promoteur de l'opéron. La fixation de ParM au promoteur réprime sa transcription.
- Localisation intracellulaire du plasmide R1 et de la protéine ParM par contraste de phase en microscopie à fluorescence. Extrait de (Moller-Jensen *et al.*, 2003).
- Filament unique de ParM polymérisant en présence d'ATP entre deux billes magnétiques recouvertes de complexe de partition ParR-*parC*. Image de microscopie électronique avec coloration négative. Extrait de (Garner *et al.*, 2004).
- Ségrégation des plasmides par des filaments de ParM de type actine. Les plasmides sont en orange ParR en bleu et ParM en vert.
- Mécanisme moléculaire de recherche et de capture.** ParM forme de courts filaments en présence d'ATP. Les polymères ont deux issues possibles : le désassemblage "catastrophe" suite à l'hydrolyse de l'ATP, ou l'élongation bipolaire suite à la fixation à chacune des extrémités d'un complexe ParR-*parC*. Adaptée de (Salje, 2010).

parRM (**Figure 17a**). Le promoteur est localisé au centre de *parC* et la liaison coopérative des protéines ParR sur *parC* autorégule la transcription de l'opéron *par* comme les protéines ParB du type **Ib** (Dam and Gerdes, 1994). Ainsi, pour tous les systèmes testés, l'opéron *par* est autorégulé, soit par l'ATPase, soit par la protéine de liaison au centromère (Jensen *et al.*, 1994; Carmelo *et al.*, 2005; Ringgaard *et al.*, 2006). Les protéines ParM sont composées d'environ 236 à 336 acides aminés (aa), alors que les protéines ParR, beaucoup plus petites sont constituées de 46 à 120 aa.

La compréhension des mécanismes impliqués dans ce système de type **II**, tient principalement de la découverte de la propriété de ParM à former des structures dynamiques, comparables aux filaments d'actine eucaryote et capables de réaliser la ségrégation des plasmides (Møller-Jensen *et al.*, 2002; Moller-Jensen *et al.*, 2003). La visualisation des premiers filaments de ParM provient d'observations par microscopie à immunofluorescence (**Figure 17b**). ParM forme des polymères dynamiques qui s'étendent le long de la cellule d'*E. coli*. Ces polymères ont pu aussi être observés *in vitro* en présence d'ATP ou de GTP (Møller-Jensen *et al.*, 2002; Popp *et al.*, 2008). Cependant, ParM se lie à l'ATP avec une affinité 100 fois supérieure au GTP, suggérant que l'ATP est le premier substrat de ParM *in vivo* (Galkin *et al.*, 2009). Les filaments de ParM présentent une forte instabilité dynamique à la fois *in vivo* et *in vitro*, résultant de l'activité d'hydrolyse de ParM (Garner *et al.*, 2004; Campbell and Mullins, 2007) (**Figure 17d**). L'utilisation d'un analogue non hydrolysable de l'ATP permet de stabiliser ces polymères (Garner *et al.*, 2004), en formant une protection contre la dépolymérisation appelée "capping" (**Figure 17e**). La présence de complexes de partition associés à chaque extrémité du filament de ParM permet également de les stabiliser. Ainsi, l'allongement ou le rétrécissement des filaments est directement régulé par cette activité d'hydrolyse de l'ATP.

Des analyses de microscopie en time-lapse ont montré que des plasmides appariés pouvaient se déplacer vers les pôles opposés de la cellule en quelques secondes seulement (Campbell and Mullins, 2007). Cette observation suggère un mécanisme dans lequel les filaments pourraient rapidement pousser les plasmides aux pôles opposés de la cellule, puis dépolymériseraient. Cet appariement de deux plasmides a été observée directement en ME pour le plasmide R1 (Jensen *et al.*, 1998). De manière remarquable, ce système de partition a pu être reconstitué *in vitro* avec seulement les trois composants du système de partition et de l'ATP (**Figure 17c**). Ceci indique que le système de partition ParRMC est suffisant pour faire la

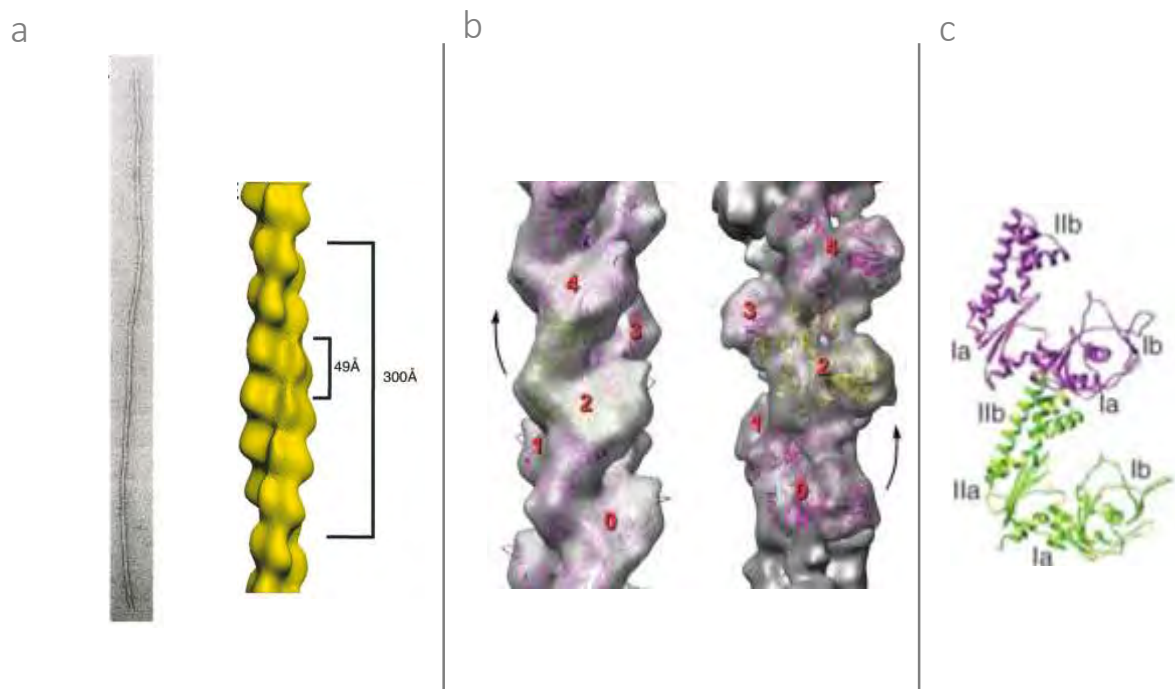


Figure 18 *Structure des filaments de ParM*

- a) A gauche un filament de ParM, visualisé en ME et à droite une reconstruction tridimensionnelle d'un filament de ParM. Extrait de (van den Ent *et al.*, 2002).
- b) A gauche, une reconstruction tridimensionnelle d'un filament de ParM avec en transparence cinq sous-unité de ParM. Les flèches représentent l'enroulement gauche du filament. Pour comparaison, une reconstitution d'un filament d'actine-F est montrée à droite. Ce filament adopte un enroulement droit. Extrait de (Orlova *et al.*, 2007).
- c) Structure d'un filament, constitué d'un enchainement de deux dimères de ParM. L'interface d'interaction la plus forte entre deux dimères dans un filament ParM, est entre le sous-domaine IIa d'une sous-unité et le sous-domaine Ia de la sous-unité du dessus. Extrait de (van den Ent *et al.*, 2002).

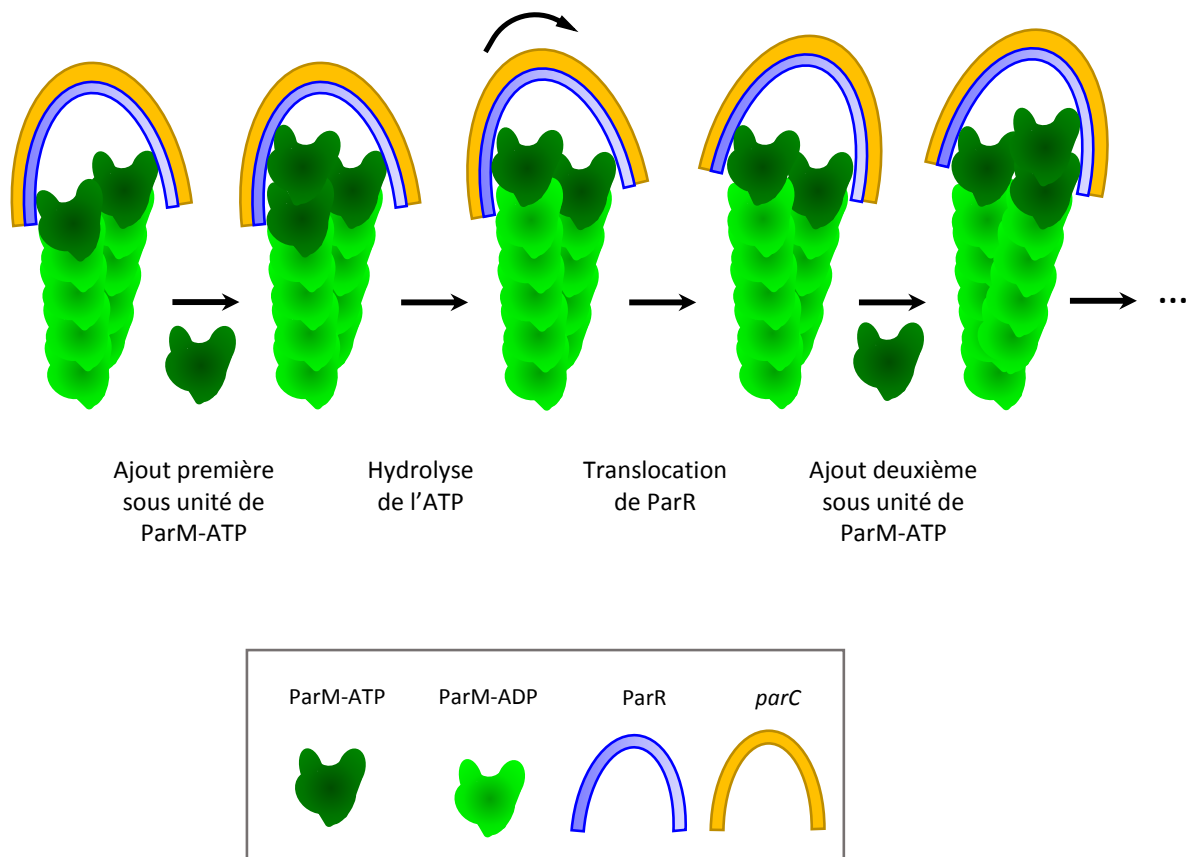


Figure 19 | Modèle de la polymérisation de ParM en "marche d'escalier"

L'hélice de ParR, fixée à *parC*, reconnaît les sous unités ParM-ATP. Une seule sous unité ParM-ATP s'ajoute sur un des deux protofilaments. L'hydrolyse de l'ATP crée une translocation processive de l'hélice ParR. Une nouvelle sous unité de ParM-ATP s'insère sur l'autre protofilament et engage un deuxième cycle.

Adaptée de (Salje and Löwe, 2008).

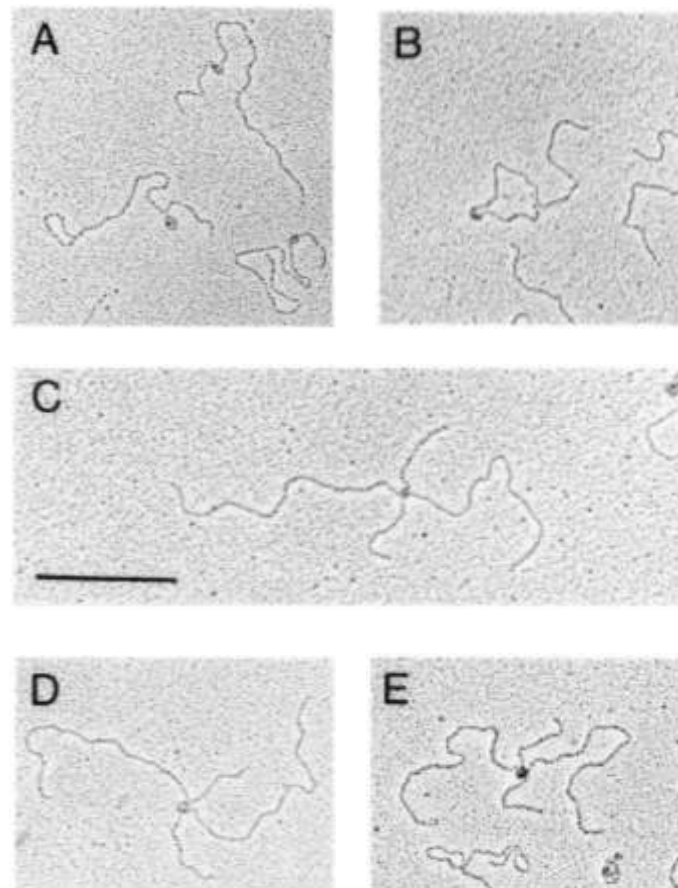


Figure 20 *Appariement des plasmides R1*

(A-E) Observation en ME des complexes de partition ParR-*parC* formés sur les plasmides R1 et de l'appariement spécifique de deux molécules d'ADN au niveau des complexes de partition. La barre noire correspond à 1 Kb.

Extrait de (Jensen *et al.*, 1998).

partition et ne nécessite pas de facteur de l'hôte ou la présence du nucléoïde pour guider la ségrégation. Dans cette reconstruction, des billes magnétiques complexées avec des sites *parC*, sont poussées en présence des protéines purifiées ParM et ParR et d'ATP (Garner *et al.*, 2004), avec une vitesse comparable à celle observée sur les analyses *in vivo* (**Figure 17c**). Des expériences de FRAP (fluorescence recovery after photobleaching) montrent que l'allongement des filaments se fait de manière bidirectionnelle, par incorporation de monomères de ParM-ATP à l'interface entre le complexe ParR-*parC* et le filament. La compréhension moléculaire de cette incorporation et de sa régulation, a nécessité deux avancées majeures : l'étude de l'architecture du complexe ParR-*parC* et l'étude des filaments de ParM. Le complexe de partition ParR-*parC* joue un rôle important au cours de ce processus. En adoptant une structure définie, il est capable de stabiliser les filaments de ParM par le phénomène de protection contre la dépolymérisation dit de "capping". Ce point sera détaillé dans le chapitre suivant.

Des images de ME et cryo-ME montrent que ParM forme des filaments ordonnés, en double hélice, similaires aux filaments d'actine à un point près. Contrairement aux filaments d'actine qui ont un pas d'hélice droit, les filaments de ParM ont un pas d'hélice gauche (Orlova *et al.*, 2007; Popp *et al.*, 2008) (**Figure 18**). Sur les images de ME, les filaments sont regroupés en "fagots" de 3 ou 5 filaments, ce qui implique de savoir si les filaments observés *in vivo* contiennent un seul ou plusieurs filaments comme le suggère ces images. Des analyses de ME utilisant un système reconstitué avec des nanocristaux d'or, montrent que deux complexes ParR-*parC* sont présents aux extrémités d'un seul et unique filament de ParM (Salje and Löwe, 2008; Choi *et al.*, 2008). L'ensemble de ces données a conduit (Salje and Löwe, 2008) à proposer un modèle en "marche d'escalier", basé sur l'ajout successif de monomères de ParM-ATP sur chaque protofilament du filament de ParM, entraînant le déplacement séquentiel du complexe ParR-*parC* (Salje and Löwe, 2008) (**Figure 19**). Ce modèle "en marche d'escalier" s'intègre parfaitement dans le modèle global des systèmes de partition de type II, dit "pushing", dans lequel les filaments de ParM s'insèrent entre les deux plasmides préalablement appariés, pour les pousser aux pôles opposés de la cellule (**Figure 20**).

De récentes analyses structurales et biochimiques indiquent qu'il existe des variations au niveau de la structure des filaments de ParM ainsi qu'au niveau de la dynamique de polymérisation des ATPases de type II. Par exemple, l'ATPase du plasmide pSK41 ne forme pas de polymère double brin comme ParM de R1, mais plutôt des polymères simple brin, bien que

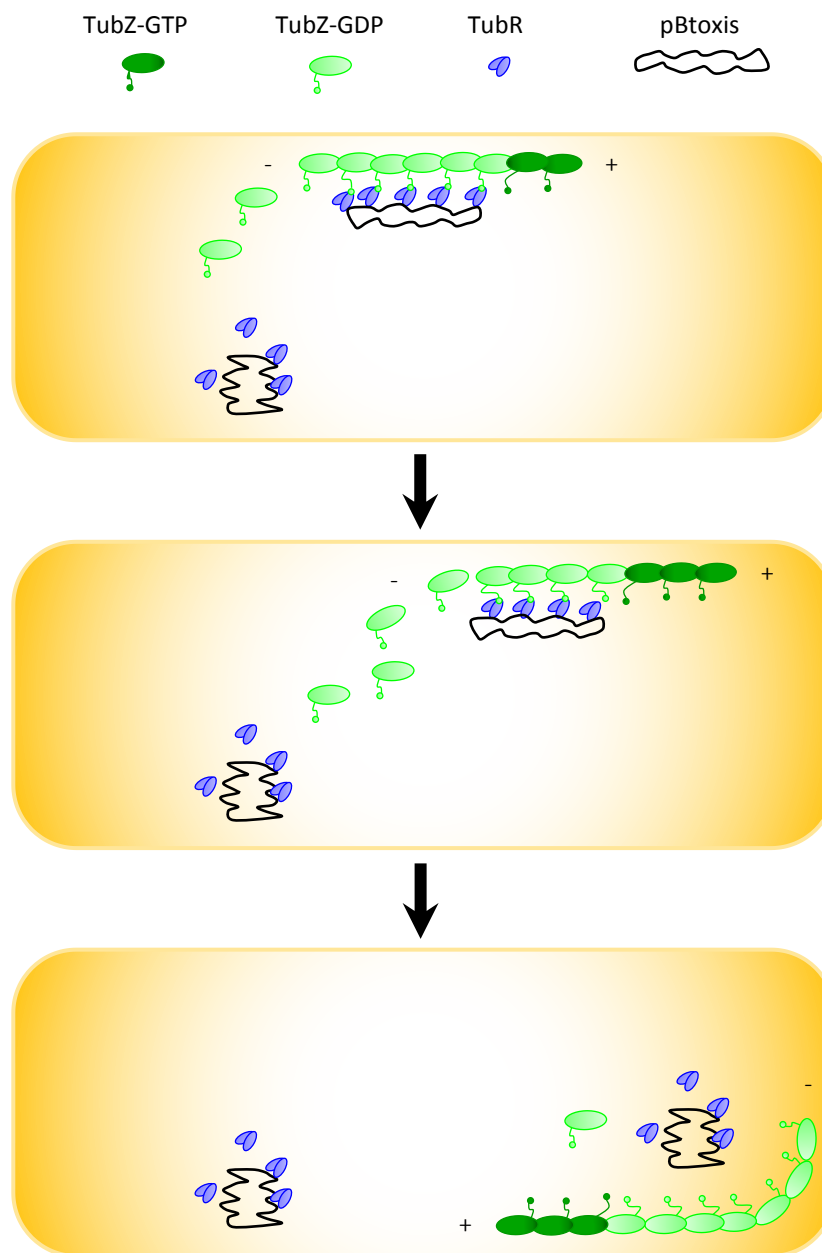


Figure 21 *Modèle en "Tapis roulant" pour les systèmes de partition de type III -Exemple du modèle de partition du plasmide pBtoxis*

Premièrement, TubR se fixe sur les sites centromériques d'une copie du plasmide pBtoxis, puis contacte la région C-terminale de la protéine TubZ (indiqué par les lignes en pointillée sur les cercle "TubZ") organisée en polymère. Par un mécanisme de "tapis roulant", les sous unités TubZ sont perdues à l'extrémité (-) du polymère et sont ajoutées à l'extrémité (+). TubR est ensuite transporté, grâce à son interaction TubR-TubZ, le long du polymère en croissance jusqu'à ce qu'il atteigne le pôle de la cellule. Le contact avec la membrane au niveau du pôle de la cellule permettrait le décrochage du plasmide. TubZ change alors de direction, contacte l'autre complexe TubR-pBtoxis et l'amène au pôle opposé de la cellule selon le même mécanisme.

la structure des deux ATPases semble similaire (D. Popp *et al.*, 2010). Les filaments formés par ParM de pSK41 sont aussi plus stables que les filaments de R1.

L'ATPase AlfA de type **II**, du plasmide pBET131 de *B. subtilis*, diffère aussi sur plusieurs points de ParM de R1 ; comme par exemple sur la structure de ces polymères qui semblent beaucoup plus enroulés que ceux de ParM de R1 et sur l'absence d'instabilité dynamique de ces filaments, *in vitro* (Polka *et al.*, 2009; David Popp *et al.*, 2010). Comme pour le type I, des variations mécanistiques existent entre les différents membres des systèmes de type **II**, notamment au niveau de la régulation et des activités de la protéine ATPase. Il n'est pas encore connu à ce jour comment ces différences influencent le processus de polymérisation des systèmes de partition de type **II**.

C. Les systèmes de partition de type III

Le système de partition de type **III** dit "tubuline" est le dernier système à avoir été découvert. Il est retrouvé sur les grands plasmides des bactéries du groupe *Bacillus cereus*, dont le plasmide de *Bacillus anthracis* (Ba) et le plasmide pBtoxis de *Bacillus thuringiensis* serovar *israelensis* (Bt) (Tang *et al.*, 2006; Larsen *et al.*, 2007; Chen and Erickson, 2008; Aylett and Löwe, 2012). Ce système, aussi appelé système TubZRC, est composé d'un adaptateur, TubR, d'une séquence centromérique *tubC* et d'une protéine TubZ appartenant à la famille des GTPases de type tubuline, principale caractéristique de ce type **III**. TubZ est capable de polymériser *in vitro* en présence de GTP et également *in vivo* (Larsen *et al.*, 2007; Chen and Erickson, 2008). Cette polymérisation est augmentée et stabilisée par la présence du complexe de partition TubRC, avec une stabilisation totale si le site centromérique est présent en entier (7 répétitions) (Oliva *et al.*, 2004; Aylett and Löwe, 2012). Cet effet de stabilisateur est aussi observé lorsque le GTP est remplacé par un analogue non hydrolysable comme le GTPγS. Le complexe de partition prévient donc la dépolymérisation des filaments de tubuline en empêchant l'hydrolyse du GTP. Ce phénomène de protection contre la dépolymérisation dit de "capping" a aussi été décrit pour les systèmes de type **II**.

L'ensemble des données recueillies, suggère un nouveau mécanisme de partition, dit "mécanisme en tapis roulant", spécifique au type **III** (**Figure 21**). Dans ce modèle, le complexe TubR-pBtoxis s'attache aux filaments TubZ par l'interaction de TubR avec les régions C-terminales de TubZ qui sont exposées en surface. L'hydrolyse du GTP au sein du polymère TubZ

génère une force en tapis roulant avec allongement des extrémités positives et rétraction aux extrémités négatives des filaments. Cette force permettrait le déplacement du plasmide préalablement fixé au polymère, vers l'un des pôles de la cellule. Après libération du complexe TubR-pBtoxis au pôle et demi-tour, le filament pourrait continuer à se déplacer vers le pôle opposé de la cellule jusqu'à la rencontre avec l'autre complexe TubR-pBtoxis. Ensuite il transporterait ce complexe vers le pôle opposé de la cellule. Ainsi, TubZ serait capable de faire des allées retours entre les pôles ce qui serait à l'origine de la séparation des plasmides. Cependant, ce système est lui aussi encore très largement incompris et plusieurs questions subsistent.

III. Le complexe de partition

La première étape de la partition implique une fixation des protéines de liaison au centromère (CBPs) ParB sur les sites centromériques, *parS*. Cette reconnaissance ParB/*parS*, permet l'assemblage d'une structure nucléoprotéique appelée le "complexe de partition". Dans les grandes familles de systèmes de partition, on observe une grande diversité au niveau des protéines ParB et du site *parS* (Hayes & Barilla, 2006), ainsi qu'au niveau de l'organisation du complexe de partition. La compréhension complète de cette organisation nécessite la combinaison de plusieurs approches comme des analyses structurales, biophysiques, combinées à des données biochimiques, génétiques et de localisation cellulaire. Dans cette partie, nous faisons le point sur l'état des connaissances sur les complexes de partition et leur rôle dans le mécanisme de partition, en rassemblant l'ensemble des études réalisées sur le sujet. La récente résolution structurale de certaines protéines ParB et de quelques complexes de partition permet d'entrevoir leur rôle dans la structuration et la mobilisation de la protéine motrice GTPase ou ATPase (ParA).

A. Le centromère nature et diversité

Le centromère ou séquence centromérique est défini comme le site d'attachement du fuseau de microtubules aux différents chromosomes eucaryotes, assurant leur ségrégation durant la mitose ou la méiose. La grande majorité des réplicons bactériens, chromosomes et

| Type | Sous-Type | Réplicon | Nomenclature des <i>parS</i> | unité centromérique minimale | nbre de répétition |
|----------|-----------|-------------------------|------------------------------|--|---|
| Type I | Ia | Plasmide P1 | <i>parS</i> | (A) ATTTCA ^A / _C (B) TCGCCA | 4 (A) 2 (B) |
| | | Plasmide F | <i>sopC</i> | TGGGACCA ^A CGGTCCA | 12 |
| | | Plasmide RK2/RK2 | <i>OB1-10</i> | TTTAGC ^G / _C GCTAAA | 10 |
| | Ib | Plasmide TP228 | <i>parH</i> | ACACTCATAAACTCATATG | 4 |
| | | pB171 (par2) | <i>parC1-2</i> | (I) ^A / _T ^T / _A CA ^T / _C A (II) TTAT ^G / _T A | 12 (I) et 5 (II) dans <i>parC1</i> , 7 (I) et 11 (II) dans <i>parC2</i> |
| | | pSM19035 | (<i>parS1-3</i>)x2 | ^A / _T ATCAC ^A / _T | 9 (<i>parS1</i>), 7 (<i>parS2</i>), 10 (<i>parS3</i>) |
| | | pTAR | <i>parS</i> | A ^C / _T CCG ^A / _G T | 13 |
| | Ic | consensus chromosomique | <i>parS</i> | tgTTNCACGTGAAACa* | 1-23° |
| Type II | | Plasmide pB171 (par1) | <i>parC1</i> | AATACTCCAAT | 2 |
| | | plasmide R1 | <i>parC</i> | AAA ^A / _C ^A / _C ^A / _C A ^A / _C ^A / _C CC | 10 |
| Type III | | pBtoxis | <i>tubC</i> | T(^T / _A)(^T / _A)(^C / _A)(^G / _A)GTTTA(^A / _C)(^A / _C) | 7 |

* Sequence consensus basée sur différents réplicons de cette catégorie

° Nombre qui varie entre les espèces

Tableau 4 *Caractéristiques et variations des séquences centromériques*

Les abréviations d'espèces pour le type Ic sont les suivantes : Bsu *Bacillus subtilis*, Tth *Thermus thermophilus*, Ccre *Caulobacter crescentus*, Ppu *Pseudomonas putida*, Pae *Pseudomonas aeruginosa*, Scoe *Streptomyces coelicolor*. (A) et (B) correspondent aux boîtes du centromère du plasmide P1. n.i: non identifié.

plasmides à faible nombre de copies, portent également des séquences centromériques, génériquement dénommé *parS*. Il existe une grande hétérogénéité de nomenclature entre les différents sites *parS* étudiés (**Tableau 4**). Bien que leur fonction soit la même, le positionnement, la taille et le nombre de séquences centromériques, peuvent être radicalement différents selon le réplicon qui les porte.

1. Les centromères chromosomiques

Les premiers sites *parS* chromosomiques identifiés, ont été décrits chez *B. subtilis* (Lin and Grossman, 1998). Par Immuno-précipitation, les auteurs ont montré que les 8 séquences *parS* situées dans la région origine, étaient capables de se lier, *in vivo*, à SpoOJ (ParB). Chaque site *parS* contient une séquence de 16pb, avec un motif palindromique de 8pb. L'utilisation de ce site consensus *parS* (5'-TGTTNCACGTGAAACA-3'), leur a également permis d'identifier des séquences centromériques homologues sur les 10 autres génomes disponibles dans la base de données de cette époque. Par une approche expérimentale, l'identification des sites *parS* chromosomiques s'est étendue à sept autres espèces bactériennes : *P. aeruginosa*, *B. cenocepacia*, *P. putida*, *S. coelicolor*, *Helicobacter pylori*, *C. crescentus* et les *Vibrio* (Mohl and Guber, 1997; Kim *et al.*, 2000; Mohl *et al.*, 2001; Godfrin-Estevenon *et al.*, 2002; Bartosik *et al.*, 2004; Dubarry *et al.*, 2006; Lee *et al.*, 2006; Yoshiharu Yamaichi *et al.*, 2007). Dans presque tous les cas, ces sites *parS* chromosomiques sont très similaires à la séquence consensus de *B. subtilis* à la fois dans leur structure, leur longueur et leur séquence. Des sites *parS* ont également été identifiés sur les chromosomes secondaires de *V. cholerae* et *B. cenocepacia*, deux espèces bactériennes aux génomes complexes. Dans ces deux organismes, les séquences *parS* du chromosome principal sont presque identiques à celle de *B. subtilis*, alors que les sites de *parS* des chromosomes secondaires diffèrent significativement (Kim *et al.*, 2000; Dubarry *et al.*, 2006).

Une approche bio-informatique a confirmé les résultats précédents et a permis d'identifier que sur 400 séquences chromosomiques disponibles dans la base de données NCBI en 2007, 276 (69%) possèdent des sites *parS* (Livny *et al.*, 2007) (**Figure 22**). Étonnamment, bien que la matrice consensus *parS* ait été calculée à partir de seulement deux espèces à Gram positif, *B. subtilis* et *S. coelicolor*, des sites *parS* conservés ont été retrouvés dans toutes les

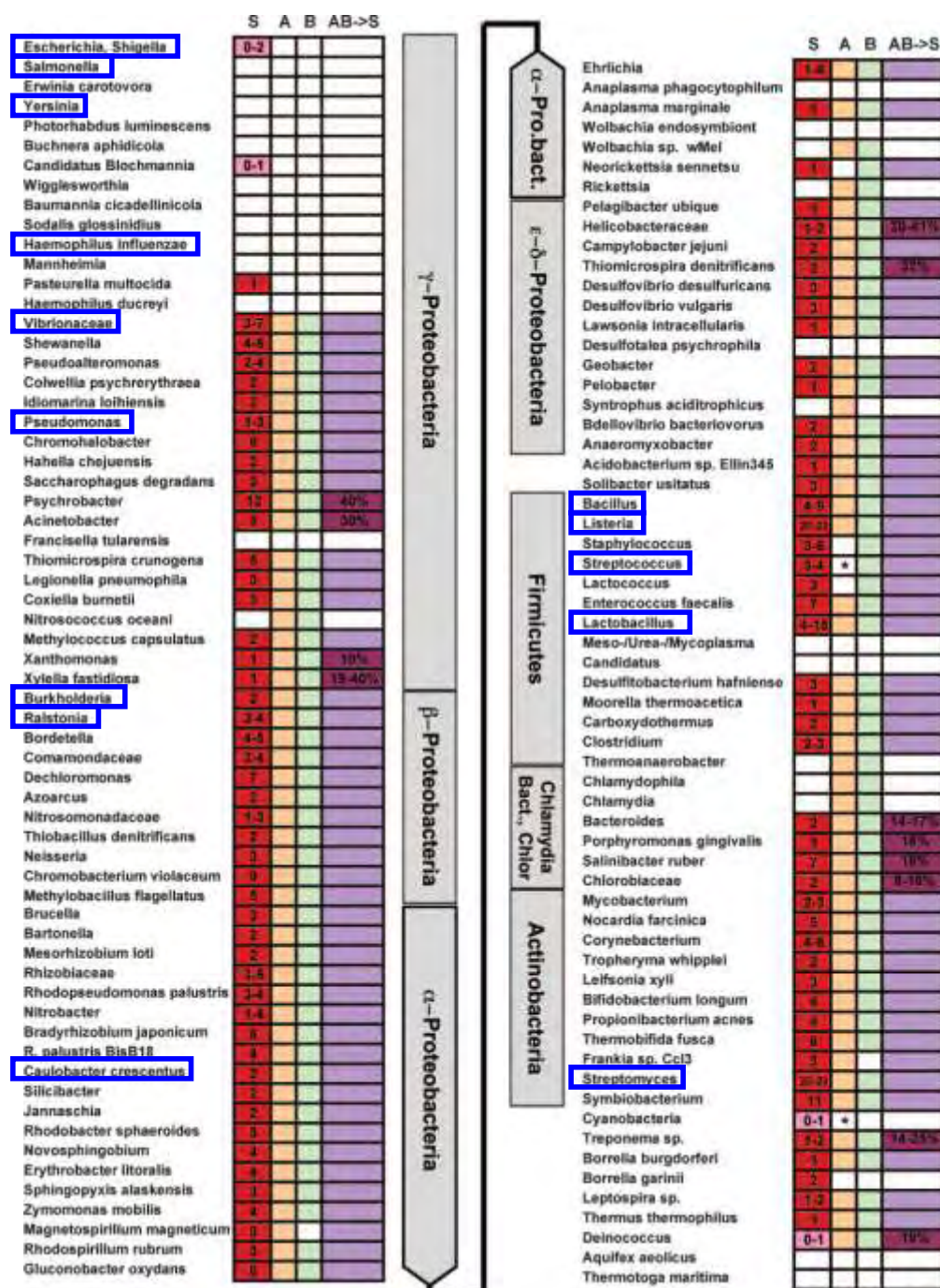


Figure 22 | Répartition des systèmes parABS sur le chromosome principal des bactéries de l'arbre phylogénétique bactérien

Les chiffres dans la colonne S correspondent au nombre de sites *parS* présents dans les différentes souches d'une branche. Les cellules rouges indiquent que toutes les souches de la branche comportent des *parS*; les roses, que seules certaines souches en comportent; les blanches, qu'aucun *parS* n'est identifiable. Dans les colonnes A et B, les cellules sont colorées si toutes les souches de la branche comportent un homologue putatif de *parA* ou *parB*, respectivement. Dans la colonne AB-S, une coloration claire indique que dans toutes les souches les trois partenaires sont distants de moins de 10% de la taille du chromosome; une coloration sombre indique que, pour au moins une souche, la distance entre *parA* ou *parB* et le site *parS* le plus proche est supérieure à 10% (les chiffres indiquent la distance dans ces souches). Les classes et phyla sont indiqués par les boîtes grisées. Bact et Chlor correspondent aux Bacteroidetes et Chlorobi. Les branches dont il est fait mention dans le texte sont encadrées en bleu.

Extrait de (Livny *et al.*, 2007).

branches de l'arbre évolutif procaryotes, y compris dans quatre souches d'archées. Deux branches de γ -protéobactéries, incluant *E. coli*, *Salmonella sp*, *Haemophilus sp*, et *Yersinia sp*, sont dépourvues de site *parS* et du système *ParABS* au complet. A l'exception de ces deux branches de γ -protéobactéries, les organismes qui n'ont pas de sites *parS* sont dispersés dans tout l'arbre évolutif procaryote (Livny *et al.*, 2007). L'absence de site *parS* est restreint à très peu de branches dans l'arbre des espèces procaryotes. Par conséquent, (Livny *et al.*, 2007)ses collaborateurs, suggèrent que les systèmes *ParABS* sont apparus très tôt dans l'évolution des chromosomes procaryotes et que l'absence de site *parS*, de *ParA*, et/ou de *ParB* dans certaines souches, reflète probablement une perte beaucoup plus tardive, d'un ou plusieurs de ces loci, dans l'évolution procaryote.

Dans la grande majorité des cas, les sites de *parS* sont relativement proches des loci *parAB*, à l'exception de *H. pylori* et sont concentrés dans les 20% du chromosome autour de l'origine de réplication et (Lin and Grossman, 1998; Lee and Grossman, 2006; Livny *et al.*, 2007). Par ailleurs, une grande variabilité du nombre de sites *parS* existe entre les différents genres bactériens ainsi qu'entre différentes espèces d'un même genre. Majoritairement, on retrouve environ 1 à 8 sites par chromosome mais ce chiffre peut être supérieur dans le cas des espèces *Listeria* (20-22), *Lactobacillus* (4-15) et *Streptomyces* (22-23) (Livny *et al.*, 2007). Dans le cas des bactéries multichromosomiques, aucun site *parS* caractéristique des chromosomes primaires, n'a été identifié sur les chromosomes secondaires ce qui est en accord avec les observations de (Dubarry *et al.*, 2006). Les séquences *parS* trouvées sur les chromosomes secondaires, *parS_{ChII}*, sont famille-spécifiques. Par exemple, la séquence consensus de *V. cholerae* du chromosome II (ChrII), n'est retrouvée que sur les ChrII de l'espèce *Vibrionaceae* / *Photobacteriaceae*, mais pas dans les espèces en dehors de cette famille. Globalement, il y a une divergence significative des séquences *parS_{ChII}* trouvées dans des familles différentes, ce qui est cohérent avec l'idée que les chromosomes secondaires ont des histoires évolutives nettement plus diversifiées que celles des chromosomes primaires, à l'image des plasmides. Contrairement au système *par* codé par le chromosome I de *V. cholerae* (Fogel and Waldor, 2006), le locus *par* codé par le chromosome II, est nécessaire à sa ségrégation comme pour les plasmides (Y. Yamaichi *et al.*, 2007). Ainsi les chromosomes primaires et secondaires d'une même espèce ne possèdent pas le même site *parS*, afin d'éviter les problèmes d'incompatibilité entre réplicons portant le même système de ségrégation, comme cela a été observé dans les

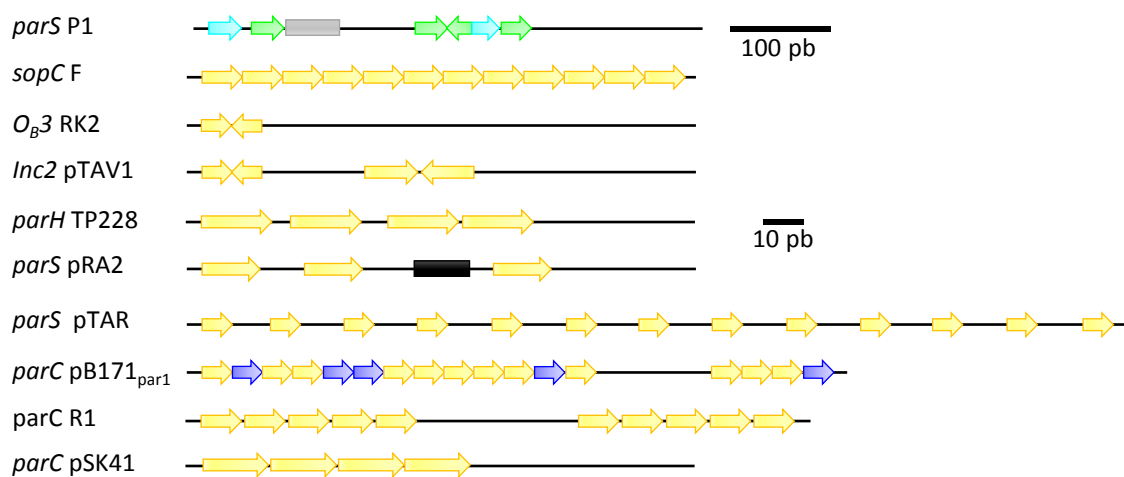


Figure 23 / Organisation des centromères plasmidiques

Ne sont représentés ici que les centromères plasmides les mieux caractérisés. Les séquences répétées sont indiquées par des flèches. Les séquences de ces répétitions varient selon les centromères. Dans *parS* de P1, les boîtes B sont indiquées par des flèches bleues, les boîtes A par des flèches vertes et le site de liaison pour IHF est représenté par un rectangle gris. La boîte noire du *parS* du plasmide pRA2 diffère par sa séquence des trois répétitions directes de cette région. Le site *parC* de pB171_{par1} est également composé de deux motifs avec des répétitions différentes, désignées par des flèches jaunes et bleues. Il faut noter la différence d'échelle entre *sopC* du plasmide F et celle des autres sites.

Adaptée de (Hayes and Barillà, 2006).

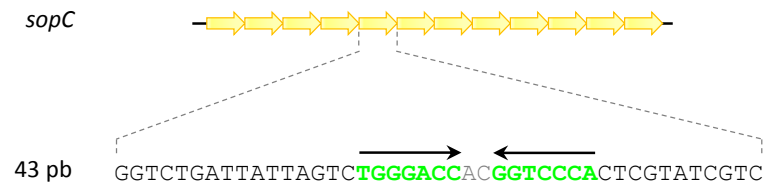
cellules hébergeant des plasmides qui contiennent des sites *parS* similaires (Bouet, Nordström, *et al.*, 2007). Contrairement aux eucaryotes qui peuvent utiliser un appareil mitotique unique pour réaliser la ségrégation de plusieurs réplicons, les procaryotes utilisent des systèmes de partition différents pour chaque réplicon d'une même cellule.

2. Les centromères plasmidiques

Les sites *parS* plasmidiques sont localisés proche des gènes *parAB*. Leur position relative vis-à-vis de ces derniers est différente et propre à l'organisation génétique de chaque système de partition (**Figure 23**). Pour les plasmides P1 et F et les autres plasmides de type Ia, le centromère est positionné en aval de l'opéron. En revanche, le centromère est situé en amont des opérons pour les autres plasmides avec des protéines ParA plus courtes. Au niveau de ces derniers opérons, la région de régulation et le centromère peuvent se chevaucher (Kalnin *et al.*, 2000; Kwong *et al.*, 2001). Les centromères plasmidiques sont généralement constitués de séquences répétées en tandem et/ou inversées avec une symétrie interne (**Figure 23**). Le nombre de répétitions ainsi que leur organisation, peut varier d'un plasmide à l'autre (**Figure 23**). Le centromère *sopC* du plasmide F est composé de 12 répétitions en tandem (Biek and Shi, 1994) et le site *parS* de pTAR d'*Agrobacterium tumefaciens*, contient 13 répétitions heptamériques qui sont séparées par un tour d'hélice à chaque fois (Gallie and Kado, 1987; Kalnin *et al.*, 2000). En revanche, le site centromérique O_B3 du plasmide RK2 est un simple palindrome de 13 pb. Ce site est essentiel et suffisant pour maintenir le plasmide, cependant 9 autres sites, relativement proches, pouvant potentiellement contribuer à la ségrégation, ont été retrouvés de manière dispersée sur ce plasmide. Le centromère *tubC* de pBtoxis de *Bacillus thuringiensis*, est constitué de 7 séquences centromériques de 12 pb chacune. Ces séquences sont organisées en deux clusters séparés de 54pb. Le premier cluster contient 3 séquences directement répétées et le deuxième en contient 4 (Tang *et al.*, 2007). Cette organisation en cluster est aussi retrouvée chez le plasmide R1 où 10 répétitions de 11pb sont réparties en deux clusters de 5 répétitions en tandem, séparées par une région de 39 pb, incluant le promoteur *parRM* (Dam and Gerdes, 1994).

Au sein d'une même famille bactérienne, le nombre de paires de bases (pb) par répétition peut aussi être très variable. Par exemple, *sopC* du plasmide F est constitué de 12 répétitions en tandem, d'une séquence de 43 paires de bases (**Figure 24a**). Cette séquence possède un site de fixation de SopB de 16 pb (Hayakawa *et al.*, 1985; Mori *et al.*, 1986). Ce

a



b

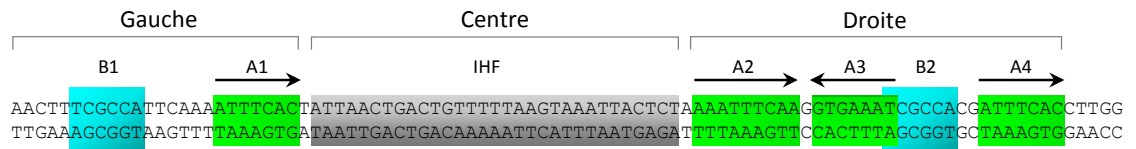


Figure 24 ∫ Sites centromériques *sopC* et *parS* du plasmide *F* et du prophage *P1*

- a) *sopC* est composé de douze répétitions d'un motif de 43 pb contenant une séquence de 16 pb avec un inversé-répété de 7 pb.
- b) Les rectangles verts et bleus représentent respectivement les motifs des boîtes A et B de *parS* de *P1*. La région de fixation d'IHF est représentée par un rectangle gris. Le sens des répétitions des boîtes A est indiqué par des flèches noires.

dernier contient un motif inversé-répété de 8 pb (TGGGAC), séparées par 2pb (AC). Une étude réalisée sur les homologues de séquence du centromère *sopC*, montre que sur ces 22 séquences centromériques, les répétitions mesurent 43pb ou dans cinq cas 45pb (Pillet *et al.*, 2011). De même, les auteurs ont pu montrer que de petites variations peuvent apparaître au sein du site de liaison de SopB entre les différentes répétitions d'un même centromère, ce qui pourrait avoir un impact sur la fixation de la protéine SopB. Ce point sera développé dans le **Chapitre I des résultats**, en précisant les bases importantes pour la fixation à SopB. Bien qu'une seule des répétitions soit suffisante pour la fixation de SopB et pour la partition (Biek and Shi, 1994), les 12 répétitions sont importantes afin que l'efficacité de partition soit optimale.

Les centromères les moins complexes sont *parH* du plasmide TP228, qui se compose de 4 répétitions de 19 pb ; et *parC* du plasmide pSK41 de type **II**, composé de 4 répétitions directes de 20 pb (Barillà and Hayes, 2003; Schumacher, Glover, *et al.*, 2007).

Le centromère le plus complexe est le centromère *parS* du plasmide P1 (**Figure 24b**). Le centromère du plasmide P7 d'*E. coli* lui est similaire (Davis *et al.*, 1990; Funnell, 1991; Funnell and Gagnier, 1993; Hayes and Austin, 1993; Surtees, 2001). De nombreux variants de ce site ont été identifiés sur quelques plasmides de bactéries à Gram-négatif et illustrent bien la variation existant même entre plasmides très proches (Ludtke *et al.*, 1989; Hayes *et al.*, 1993; Hayes and Austin, 1993; Youngren *et al.*, 2000; Sergueev *et al.*, 2005; Dabrazhynetskaya *et al.*, 2005). La séquence centromérique du plasmide P1 est composée de quatre répétitions d'une séquence d'ADN de 7 pb appelée boîte A (ATTCAC/A) et de deux répétitions d'une séquence de 6pb appelée boîte B (TCGCCA). Le site *parS* est divisé en trois principales parties : gauche, centrale et droite. La partie droite de 34 pb, interagit spécifiquement avec ParB *in vitro* au niveau de ces boîtes A₂₋₄ et B₂ (Davis *et al.*, 1990; Funnell and Gagnier, 1993). Cette partie est suffisante pour permettre la partition *in vivo* (Martin *et al.*, 1987). Des variations de séquences au niveau de ces boîtes, et des variations complémentaires au niveau des acides aminés de ParB, confère la spécificité d'interaction ParB-*parS* entre les différents plasmides proches du plasmide P1 (Hayes and Austin, 1993; Youngren *et al.*, 2000; Sergueev *et al.*, 2005; Dabrazhynetskaya *et al.*, 2005). Cette spécificité empêche les problèmes d'incompatibilité entre plasmides. La partie gauche, composée des boîtes A₁ et B₁, ne supporte pas à elle seule la partition (Martin *et al.*, 1987). La région centrale contient un site de fixation pour le facteur de l'hôte IHF (Integration host Factor) d'*E. coli* (B. E Funnell, 1988). Lors de sa fixation, IHF induit

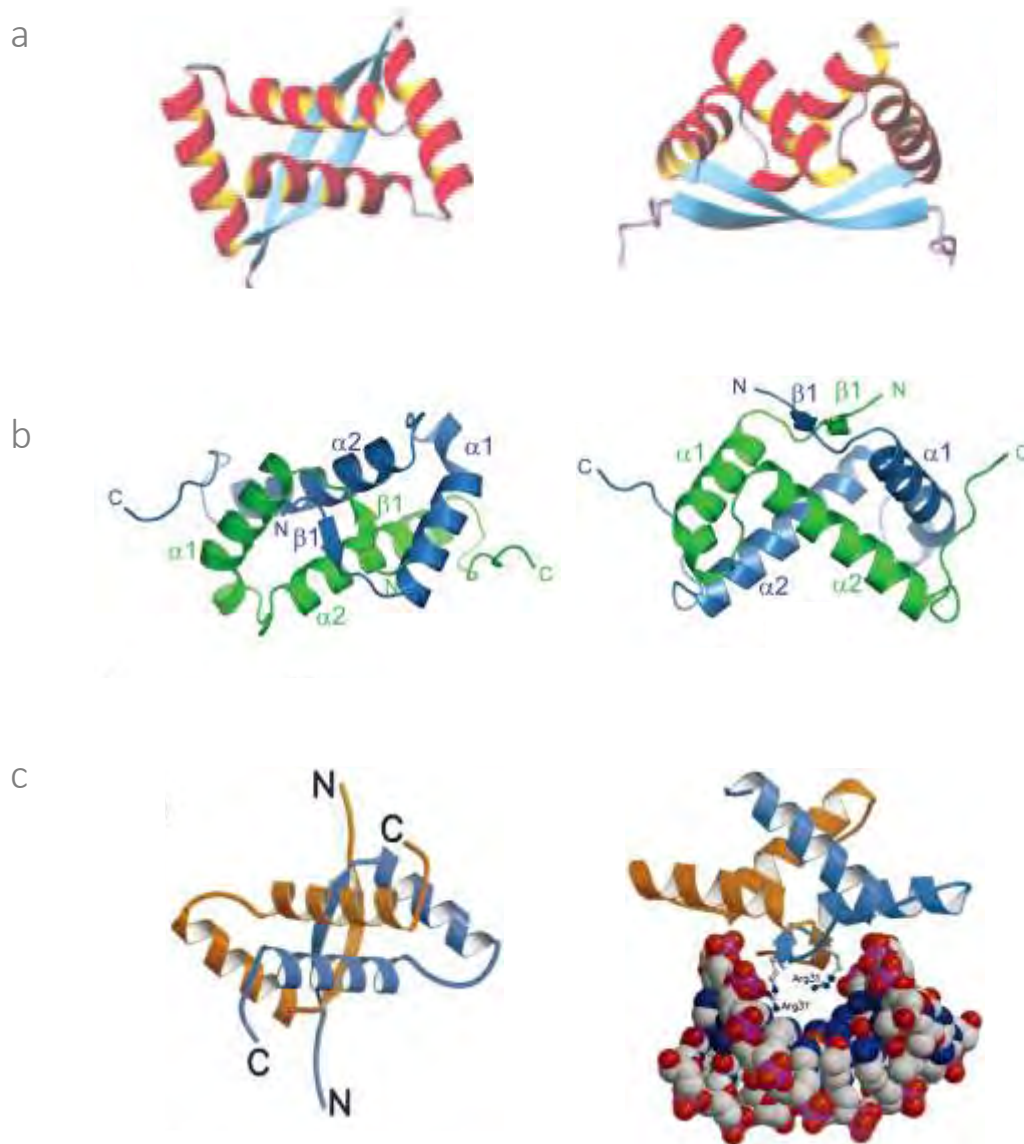


Figure 25 Structures des ParB de type Ib

- a) Représentation de deux vues du domaine RHH (résidus 34 à 74) de ParG du plasmide TP228. Cette structure a été résolue en RMN (Résonance Magnétique Nucléaire) Le domaine N-terminal de ParG, non représenté ici, est non structuré. (Golovanov *et al.*, 2003).
- b) Représentation de la structure du domaine RHH de ParB de pCXC100 selon deux vues. Les deux unités sont de couleur bleue et verte. Les éléments de structure secondaire sont étiquetés.
- c) Structure de ω (ParB) de pSM19035 avec ou sans ADN. Les sous-unités I et II d'un dimère de ω , sont en ocre et bleu, respectivement. A gauche, vue de dessus d'un dimère de ω . A droite, modèle d'un complexe entre un dimère de ω et un ADN. Les feuilletts β des deux monomères se lient au grand sillon de l'ADN. Les résidus de contacts possibles sont l'Arg31 et l'Arg310. (Murayama *et al.*, 2001).

une courbure de cette région centrale de *parS*, permettant ainsi le rapprochement des parties gauche et droite de *parS* (Funnell, 1991; Funnell and Gagnier, 1993; Rice *et al.*, 1996). Les fixations de ParB et d'IHF sont coopératives et IHF augmente l'affinité de ParB pour *parS* d'un facteur d'environ 10000 (Funnell, 1991). IHF n'est pas indispensable pour la partition de P1, toutefois son absence rend la partition moins efficace du fait de l'affinité très réduite de ParB pour *parS* (Funnell, 1991; Funnell and Gagnier, 1993).

B. Les protéines adaptatrices ou de liaison au centromère : Nature et diversité

Toutes les protéines de fixation au centromère (CBPs), sont des protéines multifonctionnelles et ont des domaines multiples qui contiennent des régions étendues flexibles, généralement au niveau de leur partie N-terminale / ou C-terminale. Les CBPs diffèrent au niveau de leur taille, celles de types **Ia** (312-343 aa) sont plus grandes que celles de type **Ib** (46-131 aa) et de type **II** (46-120 aa). En dehors du sous-type **Ia**, il est très difficile de détecter de l'homologie de séquence entre les différentes protéines CBPs. Des structures complètes ou partielles, de ces protéines, sont aujourd'hui disponibles pour chacun des types de partition : Le type **Ia**, **Ib**, **II** et plus récemment pour le type **III**. Remarquablement, la comparaison de ces structures indique que, malgré le manque significatif d'homologie de séquence entre ces protéines, elles sont toutes capables de se fixer spécifiquement à l'ADN centromérique, sous forme dimérique. Selon le motif de liaison à l'ADN qu'elles portent, deux classes de CBPs ont pu être établies : celles contenant un motif HTH sont de la famille type **Ia**, et celles contenant un motif RHH (Ribbon-Hélice-Hélice) sont de type **Ib** ou de type **II**.

Dans le cas des motifs RHH, aucune analyse des séquences primaires d'aa ne peut prédire la structure que la protéine va adopter. Il a fallu attendre la structure de ParR de pB171_{parI} de type **II** et ParG de TP228 de type **Ib** pour se rendre compte que ces deux types faisaient partie d'une même famille structurale avec le motif RHH caractéristique (Golovanov *et al.*, 2003; Møller-Jensen *et al.*, 2007). Pour les systèmes de partition de type **Ib**, trois structures des protéines adaptatrices sont aujourd'hui disponibles : ParG de TP228, ω de pSM19035 de *Streptococcus pyogenes* et ParB du plasmide pCXC100 d'une sous espèce de *Leifsonia xyli, cynodontis* (**Figure 25**) (Murayama *et al.*, 2001; Golovanov *et al.*, 2003; Huang *et al.*, 2011). Elles contiennent toutes un motif RHH, ce qui suggère fortement que toutes les protéines adaptatrices de type **Ib** contiennent ce motif de liaison à l'ADN. La seule protéine

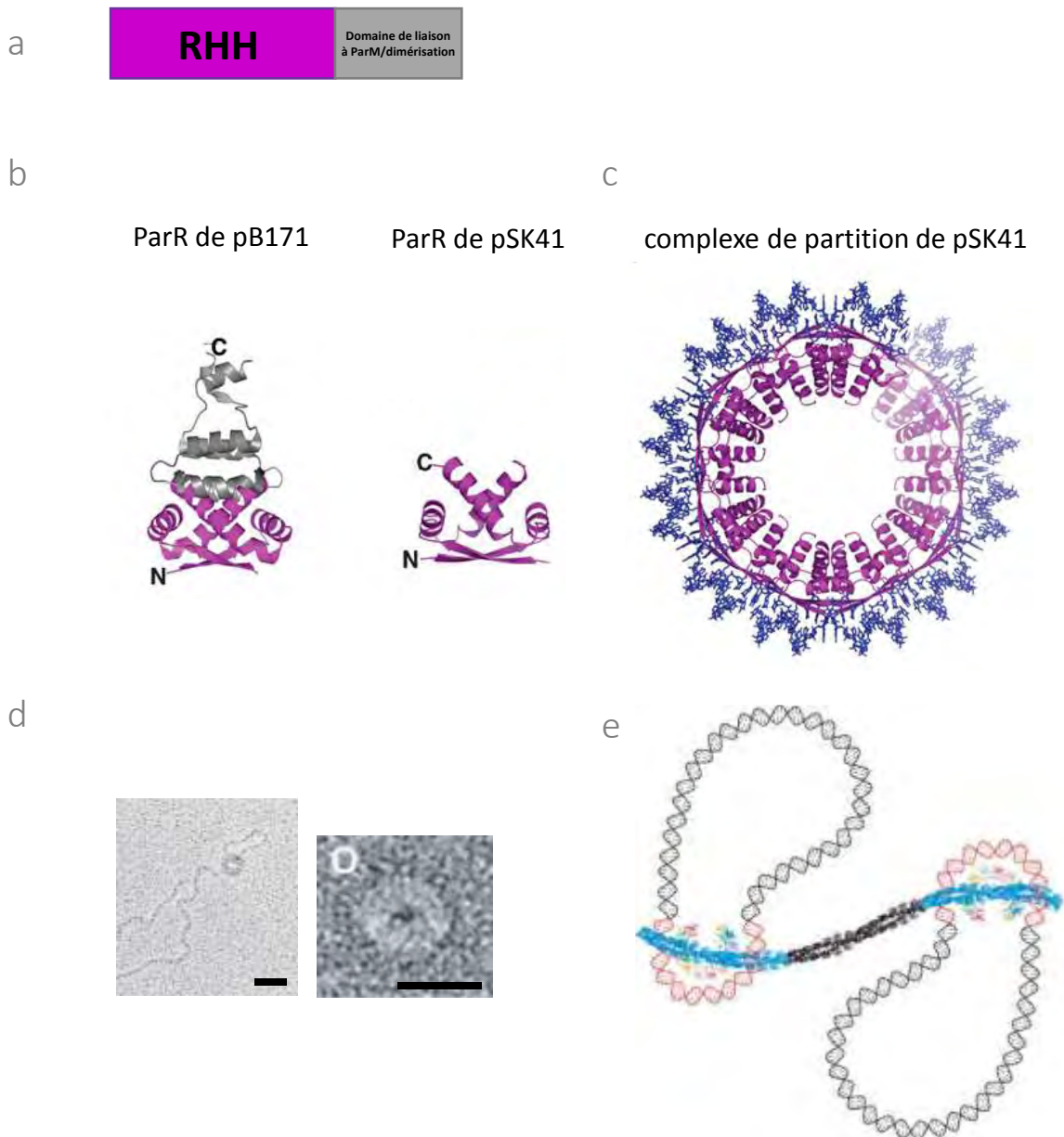


Figure 26 | Structures des protéines ParR et des complexes de partition de type II

- Organisation des domaines des protéines ParR.** Le domaine N-terminal (magenta) forme un motif RHH, responsable de la liaison au centromère. La région C-terminale est un domaine partiellement flexible qui contribue à la liaison avec ParM et à la stabilisation du dimère des motifs RHH et du dimère-dimère. Les motifs RHH des structures (b) et (c) sont représentés dans les mêmes orientations afin de pouvoir les comparer.
- Structure de ParR de pB171 d'*E. coli*.** Le domaine N-terminal RHH est de couleur magenta, et la région C-terminale, qui se compose de trois hélices et les résidus supplémentaires flexibles, sont en gris. Les régions N et C-terminales d'un monomère de ParR sont marquées (N) et (C). Extrait de (Møller-Jensen *et al.*, 2007).
- Structure de ParR du plasmide pSK41 de *S. aureus*.** Sur la gauche est représentée la structure de la protéine en l'absence d'ADN, et sur la droite se trouve le complexe ParR-centromère. L'ADN est affiché en bleu. Extrait de (Schumacher, 2008).
- Image du complexe de partition ParR-parC du plasmide R1 en ME après coloration négative. Les barres noires correspondent à 25 nm. Extrait de (Møller-Jensen *et al.*, 2007).
- Représentation schématisée de l'interaction entre ParR-parC et les deux extrémités d'un filament de ParM.** Les domaines N-terminaux de 12 dimères de ParR se fixent à parC (hélice ADN rouge), et les domaines C-terminaux interagissent avec l'extrémité des filaments composée de ParM-ATP (bleu). Extrait de (Møller-Jensen *et al.*, 2007).

adaptatrice de type **Ib** qui a été résolue liée à de l'ADN, est ω (**Figure 25c**). L'ADN utilisé dans cette étude est un ADN non spécifique donc les auteurs n'ont pas résolu comment ω se lie aux répétitions directes et inversées de son centromère (Weihofen *et al.*, 2006). Ainsi, la structure adoptée par les complexes de partition **Ib** reste inconnue.

C. Structure et formation des complexes de partition

Malgré l'apparente diversité entre les différents centromères et les protéines de liaison au centromère, l'interaction entre les deux conduit dans tous les cas à la formation d'un complexe de partition d'ordre supérieur avec une architecture particulière. Cette architecture semble être un prérequis important dans la capture et l'activation de la protéine motrice responsable de la ségrégation. Comme on le verra dans cette partie, la structure récente du complexe ParR-centromère du plasmide pSK41 de type **II**, montre une superstructure en U, aux dimensions idéales pour permettre la liaison avec le filament ParM (Schumacher, Glover, *et al.*, 2007). Les différents complexes et leurs structures seront détaillés selon le type de système de partition, dans cette section.

1. Le complexe de partition de type **II**

Les systèmes de partition de type **II** sont caractérisés par une ATPase de type actine. Les protéines CBPs de type **II** appelées ParR, sont beaucoup plus petites que les protéines ParB de type **Ia** avec une taille comprise entre 46 à 120 aa. Le système de partition le plus étudié est celui du plasmide R1 d'*E. coli*. ParR_{R1} est une protéine de 13.3 KDa, capable de se fixer spécifiquement à son centromère *parC*. Ce dernier est constitué de deux clusters de cinq répétitions de 11 pb (Dam and Gerdes, 1994; Breüner *et al.*, 1996). Les 10 répétitions sont requises pour une partition efficace.

Comme pour le plasmide R1, la protéine ParR du plasmide pB171_{par1} de type **II**, se fixe sur son centromère. La structure des résidus 6 à 95 de ParR_{pB171}, a été résolue, alors que celle des résidus 1 à 5 et les 35 derniers aa en C-terminal, ne l'est pas (Møller-Jensen *et al.*, 2007) (**Figure 26**). Cette structure contient deux monomères formant une unité asymétrique. D'un point de vue topologique, chaque monomère de ParR est composé d'1 feuillet β et de 5 hélices α . Les hélices α_1 et α_2 de chaque monomère s'entrecroisent et les feuillets β de deux

monomères s'associent de façon antiparallèle pour former un motif (RHH)₂ hydrophobe. En s'insérant dans le grand sillon de l'ADN, ce motif hydrophobe permet la fixation spécifique de ParR au centromère. Les trois hélices α (α 3- α 5) dans la partie C-terminale semblent renforcer la dimérisation des motifs RHH de chaque monomère et stabilisent les interactions dimère-dimère. Comme ParR_{pB171}, la structure de ParR_{pSK41} contient un motif (RHH)₂ de fixation à l'ADN (**Figure 26b**) (Schumacher, Glover, *et al.*, 2007).

Selon la structure cristalline, ParR_{pB171} pourrait s'assembler en une structure hélicoïdale continue. Dans le cristal de ParR, les dimères sont positionnés de sorte que leur région N-terminale est orientée vers l'extérieur de la structure et leur région C-terminale vers l'intérieur (Møller-Jensen *et al.*, 2007). Ainsi, les domaines de reconnaissance de l'ADN sont orientés vers l'extérieur de la structure, ce qui pourrait permettre à l'ADN de s'enrouler autour. Le complexe nucléoprotéique de partition adopterait donc une conformation en anneau. (Ringgaard *et al.*, 2006) et collaborateurs, avaient montré que la fixation d'un dimère de ParR_{pB171} tous les 31pb, qui est l'espacement existant entre deux sites *parC* consécutifs (3 tours de double hélice), pouvait permettre la fixation des 12 dimères de ParR_{pB171} sur une même face de l'ADN, ce qui va dans le sens de cette hypothèse (Ringgaard *et al.*, 2006). Les études de l'assemblage de ParR_{R1} et de ParR_{pB171} sur leurs centromères respectifs réalisées en microscopie électronique (ME) ont confirmé que ces complexes de partition forment des structures en anneaux (Jensen *et al.*, 1998; Møller-Jensen *et al.*, 2007) (**Figure 26d**). Pour ParR_{pB171}, le diamètre de l'anneau mesure entre 15 à 20 nm ce qui est compatible avec celui décrit dans la structure. Pour les cristallographes, les anneaux observés en ME, sont révélateurs de la nature hélicoïdale des complexes de partition (**Figure 26**).

La preuve directe d'une organisation des complexes de partition en superhélice a été apportée par la cristallisation du complexe ParR-*parC* du plasmide pSK41 de *Staphylococcus aureus* (Schumacher, Glover, *et al.*, 2007) (**Figure 26c**). ParR se fixe de manière coopérative sur les dix répétitions directes de *parC* pour former un complexe nucléoprotéique. Dans le cristal, ce complexe se traduit en une structure en hélice continue contenant six dimères de dimères avec les domaines N-terminaux d'interaction avec l'ADN centromérique, orientés vers l'extérieur. Ainsi, l'ADN centromérique s'enroule autour des ParR pour former une structure en superhélice. Cette structure de 18 nm de diamètre et de nature hélicoïdale, confirme les hypothèses émises sur les complexes ParR-centromère de R1 et de pB171.

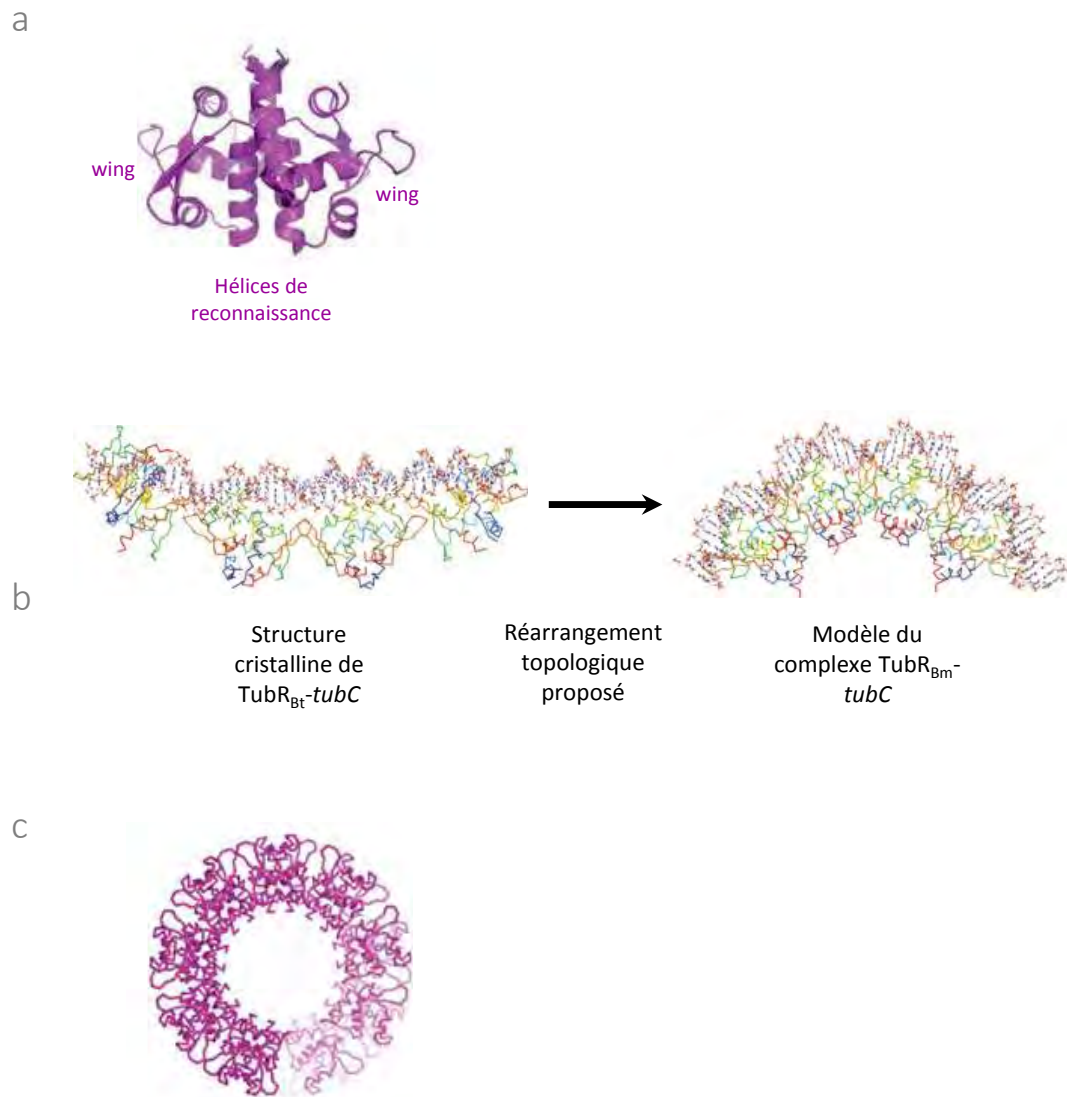


Figure 27 Structures des protéines *TubR* et des complexes de partition de type **III**

- a) Structure d'un dimère de TubR de pBtoxis. Extrait de (Ni *et al.*, 2010).
- b) **Complexe de partition des systèmes *TubZRC***. A droite, structure de TubR de pBtoxis de Bt lié à *tubC*. A gauche, reconstitution de la structure de TubR_{Bm} avec l'ADN. Extrait de (Aylett and Löwe, 2012).
- c) Reconstitution de la superstructure des protéines TubR. Extrait de (Aylett and Löwe, 2012).

Le haut degré de conservation dans l'assemblage des complexes de partition des plasmides R1, pB171 et pSK41, suggère fortement que les systèmes de partition de type **II** ont un mécanisme similaire de formation de leur complexe de partition. La superstructure hautement conservée des complexes de partition pourrait être cruciale pour la partition des plasmides, notamment pour son interaction avec la protéine ATPase ParM. Des expériences de lumière polarisée ont montré que la protéine ParR de pB171 se lie à ParM via son domaine C-terminal (Schumacher, Glover, *et al.*, 2007). Dans la structure correspondante, les régions C-terminales des protéines ParR sont dirigées vers la lumière du pore. A cet endroit, le filament de ParM pourrait s'insérer en interagissant directement avec ParR (**Figure 26e**).

L'ensemble des analyses structurales et biochimiques des complexes de partition ont permis de préciser le mécanisme de partition pour les systèmes de type **II** : ParR-*parC* se structure en une hélice ouverte en forme de "U" générant une sorte de couvercle dit "cap" qui interagit à l'extrémité supérieure d'un filament ParM, et lui permet de le stabiliser (Schumacher, Glover, *et al.*, 2007; Salje and Löwe, 2008) (**Figure 19**).

2. Le complexe de partition de type III

Les systèmes de partition de type **III** sont caractérisés par une GTPase de type tubuline. La protéine CBP de type **III** du plasmide pBtoxis de *Bacillus thuringiensis*, appelée TubR_{Bt}, n'a aucune homologie avec une protéine connue. Elle mesure 11.6 kDa. La structure d'un monomère de TubR est $\beta 1-\alpha 1-\alpha 2-\alpha 3-\alpha 4-\beta 2-\beta 3-\alpha 5$. TubR_{Bt} possède un motif boucle-HTH trouvé dans un grand nombre de protéines de liaison à l'ADN, procaryotes ou eucaryotes (Gajiwala and Burley, 2000). La dimérisation des protéines TubR se fait via deux domaines : les régions N-terminales de $\alpha 1$ et $\alpha 5$ et surtout, par un double contact entre les "hélices de reconnaissance" $\alpha 4$, de chacun des monomères (**Figure 27**). Bien que TubR montre des fortes similarités structurales avec des régulateurs transcriptionnels de la famille ArsR comme la protéine CzrA de *Staphylococcus aureus* (Eicken *et al.*, 2003; Pennella and Giedroc, 2005), l'agencement des dimères TubR entre eux est très différent, rendant ce mode de conformation singulier. Le motif HTH de TubR est formé par les hélices $\alpha 3-\alpha 4$, comme l'ensemble des ParB de type I, mais il n'est pas engagé dans une interaction HTH-ADN classique. L'hélice de reconnaissance, normalement impliquée et essentielle à l'interaction avec l'ADN centromérique, est utilisée ici pour faire la dimérisation des protéines TubR, ce qui la rend indisponible pour son interaction

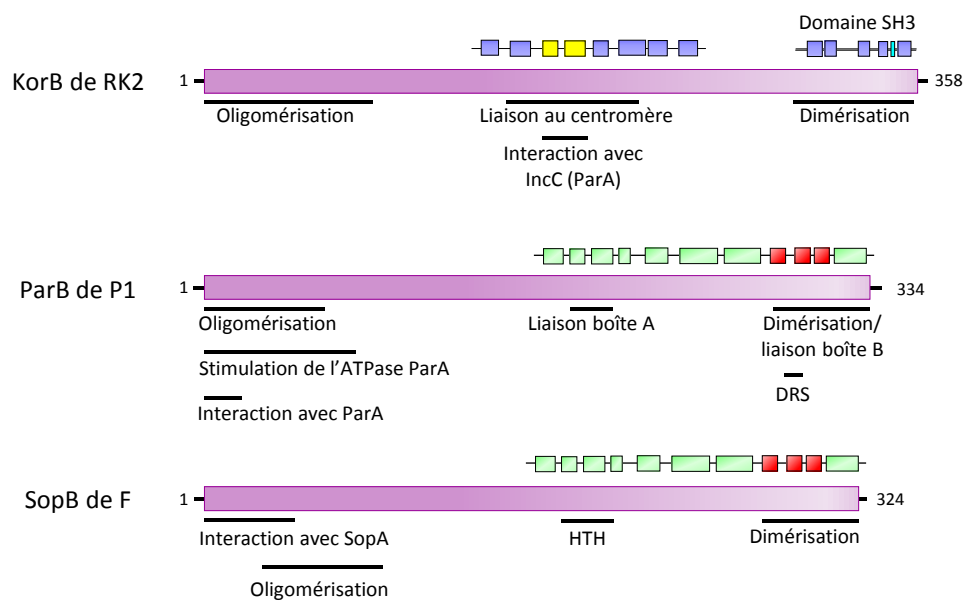


Figure 28 *Organisation des domaines fonctionnels des ParB homologues codés par le plasmide F, P1 et RK2*

Représentation des structures secondaires de SopB, ParB et KorB. Les traits noirs représentent les domaines d'interaction protéine/protéine ou les interactions protéine-ADN, cartographiées par des études de génétiques antérieures. Adaptée de (Hayes and Barillà, 2006).

avec l'ADN. Cependant, les hélices de reconnaissance de deux monomères différents créent, en se dimérisant, une interface d'interaction avec le grand sillon de l'ADN. La longueur totale de l'ADN contacté par un seul dimère de TubR_{Bt} est de 16 pb alors que la longueur de la répétition *tubC* est seulement de 12 pb (Ni *et al.*, 2010). Des études biochimiques couplées à de la modélisation structurale ont démontré que deux dimères de TubR pouvaient s'apparier au niveau de la boucle présente entre les feuillettes $\beta 2$ - $\beta 3$, appelée "wing", créant ainsi une surface d'interaction avec le petit sillon de l'ADN (au niveau des 4pb restantes). Ainsi, TubR emploie un nouveau mode de liaison à l'ADN avec une reconnaissance du grand sillon de l'ADN par les extrémités des hélices de reconnaissance d'un dimère et du petit sillon par les boucles de deux dimères différents.

En utilisant des interactions protéine-protéine multiples, TubR pourrait courber l'ADN et s'organiser en une superstructure. Dans ce sens, des images de ME montrent que TubR_{Bt} comme son homologue TubR de *Bacillus megaterium*, TubR_{Bm}, sont capables de former des structures courbées lorsqu'ils sont complexés à leur centromère (Aylett and Löwe, 2012). De même, la structure de TubR_{Bt}-centromère montre que TubR_{Bt} est capable de se fixer sur l'ADN pour former un filament linéaire légèrement courbé en hélice droite, ce qui est compatible avec un enroulement de TubR Bt autour de l'ADN. Par ailleurs, la structure en anneau adoptée par TubR_{Bm} (**Figure 27c**), rappelle fortement celle de la superstructure du complexe ParR-*parC* de R1 dans lequel l'ADN centromérique s'enroule autour des ParR. Qui de l'ADN ou de la protéine TubR s'enroule sur l'autre, est encore un sujet de débat. Quoiqu'il en soit, les complexes de partition de type **III** adoptent bien une superstructure qui est importante pour l'activité de polymérisation de TubZ et leur stabilisation (Oliva *et al.*, 2004; Aylett and Löwe, 2012).

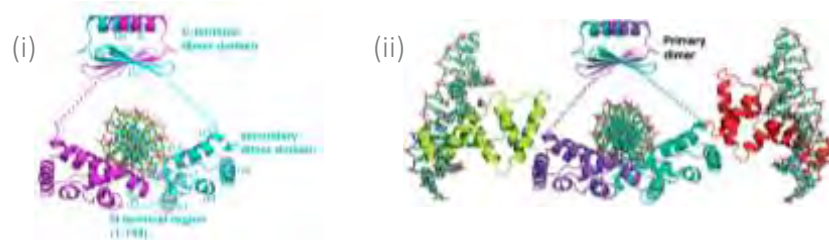
3. Les complexes de partition de type I

Les protéines adaptatrices de type **Ia** sont structurellement distinctes de leurs homologues de type **Ib** et beaucoup plus complexes. Elles sont généralement constituées de trois domaines: un domaine N-terminal de liaison à l'ATPase, un domaine central de liaison au centromère et un domaine C-terminal responsable de la dimérisation de la protéine (**Figure 28**). La comparaison de leurs séquences peptidiques indique qu'elles ont trois motifs très conservés: le motif B ne correspondant à aucun motif structural connu, le motif HTH responsable de la fixation au centromère et un motif secondaire de fixation au centromère.

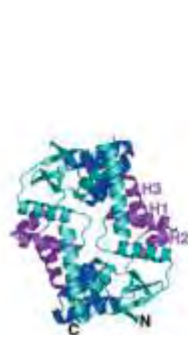
a



b



c



d



e

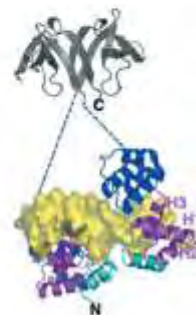


Figure 29 Structure d'un dimère de ParB (142-324) fixé ou non sur son centromère *parS*

- a) Schéma représentant l'organisation des différents domaines des protéines de liaison au centromère et le code couleur associé qui sera conservé en (c), (d) et (e). Le domaine N-terminal (en gris) est une région flexible impliquée dans l'activation et le recrutement de l'ATPase. Le motif B en cyan, fonctionne comme un domaine d'interaction dimère-dimère ou domaine d'oligomérisation. Le domaine intitulé HTH (magenta), est une unité à trois hélices comportant un motif HTH qui est très conservé dans les structures de type Ia résolus à ce jour. Le domaine dit "linker" (bleu) est un domaine de liaison entre domaine constitué d'une hélice qui relie l'unité HTH au domaine de dimérisation. Le domaine de dimérisation (gris clair) est impliqué dans la dimérisation de deux monomères de ParB. Les unités HTH de chaque structure sont présentées dans le même sens avec les trois hélices marquées H1-H3 (magenta).
- b) Structure de la protéine SopB du plasmide F liée au centromère (Schumacher *et al.*, 2010). (i) représentation d'un dimère lié au centromère. Chaque unité du centromère est colorée en bleu ou en magenta. Les domaines de dimérisation et de liaison au centromère ont été obtenus séparément ainsi cette structure de SopB a été modélisée en reliant les deux domaines par une région flexible (indiquée en pointillé). En (ii) est représentée l'interaction secondaire existant entre un premier dimère fixé sur le centromère et deux autres dimères également lié à l'ADN. Ces contacts entre dimères permettent à SopB de se lier à différentes molécules d'ADN.
- c) Structure de la protéine SpoOJ de *T. thermophilus* (Leonard *et al.*, 2004). Le domaine N-terminal n'est pas visible, bien qu'il soit présent dans la structure. Le domaine de dimérisation n'est pas présent dans la structure cependant le motif B (cyan) réalise la dimérisation de SpoOJ à des concentrations élevées. Ce motif est considéré comme un motif secondaire de dimérisation ou appelé oligomérisation. Les extrémités N-et C-terminales sont précisées sur les structures.
- d) Structure du complexe de partition ParB-(142-333)-*parS* du plasmide P1 (Schumacher and Funnell, 2005). Le code couleur est le même qu'en (a) L'HTH est représenté dans la même orientation que celle de SpoOJ et de couleur magenta.
- e) Structure de la KorB du plasmide (Delbrück *et al.*, 2002; Khare *et al.*, 2004). Le domaine de dimérisation en C-terminal (résidus 297 à 358) et le domaine de liaison d'ADN (résidus 101 à 294) ont été résolus séparément (pas tous les résidus étaient visibles) [118,119]. Le dimère de KorB a été modélisé en reliant les deux domaines par une région flexible (indiquée en pointillé). Une partie du motif B présents dans la structure est de couleur cyan. Le domaine dimère, résolu séparément, est gris et est présenté comme lié par des résidus flexibles au domaine de liaison à l'ADN. (c) et (e) ont été modélisés à l'aide PyMOL (DeLano scientifique). L'ADN est représenté en jaune.

Le motif B de fonction inconnue, est situé à proximité de la région N-terminale (**Figure 29a**). La structure de ce motif a seulement été obtenue pour la protéine chromosomique SpoOJ de *T.thermophilus*, et est constituée de 4 hélices α et de deux brins β avec une topologie $\alpha 1$ - $\alpha 2$ - $\alpha 3$ - $\beta 1$ - $\beta 2$ - $\alpha 4$ (Leonard *et al.*, 2004) (**Figure 29b et c**). Le domaine C terminal est impliqué dans la dimérisation des protéines adaptatrices au niveau de leurs 70 derniers résidus. Bien que ce domaine de dimérisation adopte des structures similaires entre les différents membres de type **Ia**, il n'est pas toujours uniquement dédié à la dimérisation. Par exemple, dans le cas de ParB du plasmide P1, ce domaine est aussi impliqué dans la liaison à l'ADN.

Quatre structures partielles de protéines adaptatrices de type **Ia** sont aujourd'hui disponibles : ParB de P1, SopB de F, KorB du plasmide RK2 et SpoOJ_{T.th} (Delbrück *et al.*, 2002; Khare *et al.*, 2004; Leonard *et al.*, 2004; Schumacher and Funnell, 2005; Schumacher, Mansoor, *et al.*, 2007; Schumacher *et al.*, 2010) (**Figure 28 et Figure 29a**). Trois d'entre elles ont été cristallisées en interaction avec de l'ADN centromérique. Cependant, aucune de ces protéines n'est sous sa forme pleine taille. En effet, les études structurales récentes sur ces protéines, ont révélé des informations importantes sur leurs fonctions de liaison de l'ADN et leur domaine de dimérisation, mais soulignent aussi les difficultés à obtenir des structures pleines tailles de ces protéines (Schumacher *et al.*, 2010). Par exemple, seuls les résidus 157 à 270 et 275-323 pour SopB, 147-333 de ParB_{P1} et les résidus 23 à 214 de SpoOJ_{T.th} ont pu être observés. Pour SopB et ParB_{P1}, les structures ont révélé la conformation du motif HTH et du domaine de dimérisation alors que pour SpoOJ_{T.th}, le motif HTH est absent de la structure. L'interaction stable du motif HTH et du centromère ou celle des deux monomères entre eux, semble permettre une meilleure visualisation de ces domaines dans le cristal. Au contraire, la partie N-terminale qui fait défaut dans ces structures semble être très flexible, comme en témoigne la sensibilité de cette partie à la protéolyse (Surtees and Funnell, 1999). Ceci pourrait expliquer la difficulté à obtenir des cristaux analysables.

La protéine SopB sert ici d'exemple pour comprendre la structure des protéines de type **I**. Cependant les similitudes et les variations des structures des autres protéines adaptatrices de type **Ia** seront discutées en parallèle. Des études génétiques et biochimiques ont mis en évidence trois domaines d'interaction pour SopB de F (Hanai *et al.*, 1996; Ravin *et al.*, 2003) (**Figure 29b**). Les résidus 1 à 42 contiennent le domaine d'interaction avec SopA, les résidus 174 à 201, le motif HTH et les résidus 245 à 323, le domaine de dimérisation. Les résidus 36 à 72

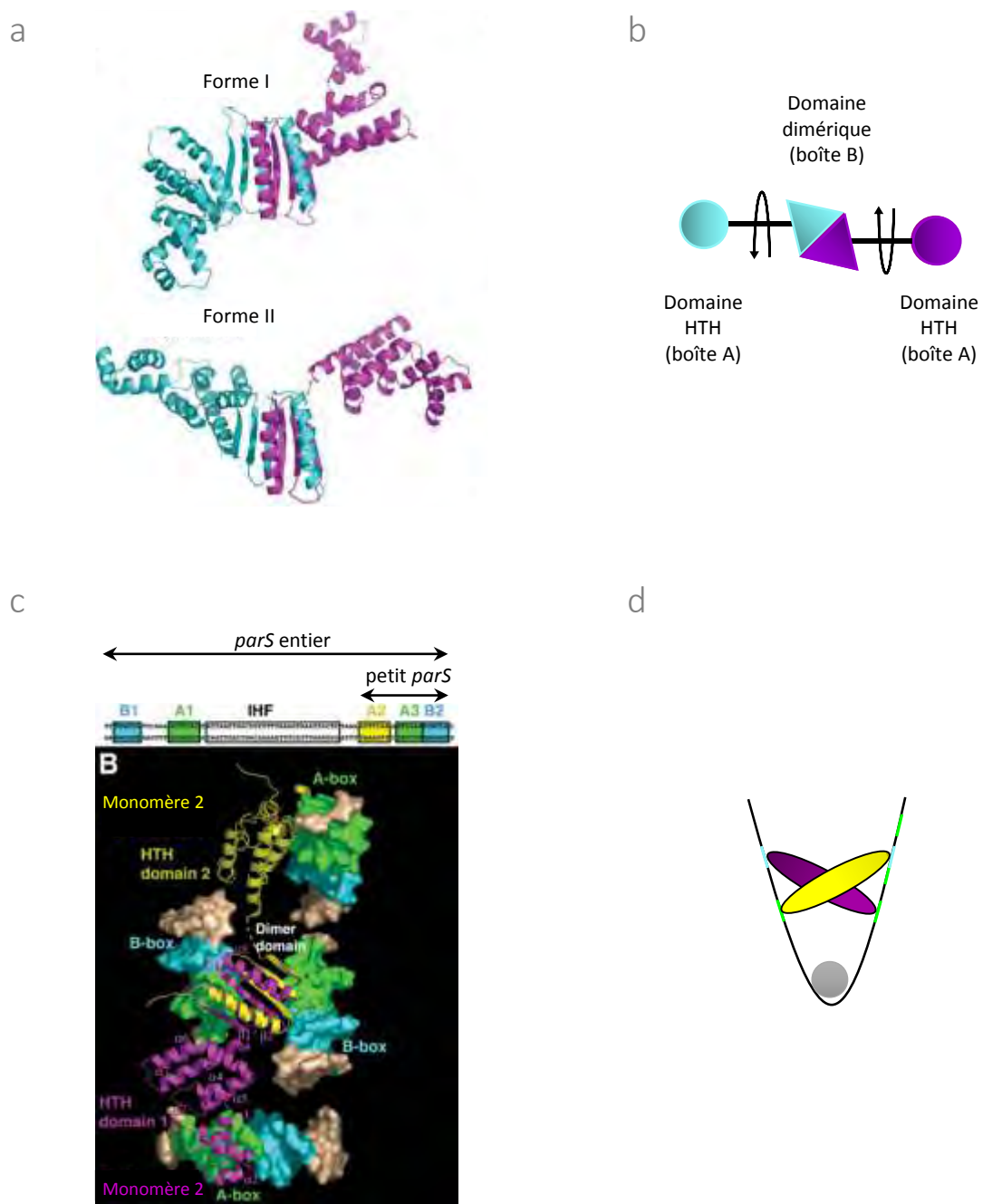


Figure 30 *Structure d'un dimère de ParB (142-324) fixé ou non sur son centromère parS*

- Structure d'un dimère de ParB avec un monomère en bleu et l'autre en magenta. Grâce à son domaine flexible entre le domaine HTH et le domaine de dimérisation, le dimère peut adopter des structures variables. (deux formes présentées ici). Adaptée de (Schumacher and Funnell, 2005).
- Représentation de la flexibilité d'un dimère de ParB (coloré comme en a). Le domaine HTH (cercle) est attaché au domaine de dimérisation (triangle) par un lien flexible. Adaptée de (Schumacher and Funnell, 2005).
- Structure d'un dimère de ParB (jaune et violet) fixé à une petite séquence *parS* composé des boîtes A3 (verte) et B2 (bleue). Le domaine HTH de chaque monomère reconnaît une boîte A3 et le domaine de dimérisation peut fixer deux boîtes B2 appartenant à deux sites *parS* différents. Extrait de (Schumacher, 2008).
- Représentation schématique de la reconnaissance des boîtes A (vertes) et B (bleues) d'un site *parS* par un dimère de ParB (bleu-magenta). Cette reconnaissance est facilitée par la courbure du site *parS*, induite par la fixation d'IHF (violet).

pourraient contenir un domaine impliqué dans l'interaction entre deux dimères de SopB, appelé domaine d'oligomérisation de SopB. En accord avec ces études et les données structurales, SopB semble être composé de trois régions (comprenant les résidus 1-156, 157-270 et 271-323). Dans le cristal, les résidus 1-156 n'ont pu être visualisés. Des analyses de dichroïsme circulaire montrent que les parties non cristallisées de SopB contiennent 23% d'hélice α et 13% de feuillet β , indiquant que ces régions sont tout de même structurées et repliées. Cependant, elles doivent être attachées de manière flexible au domaine central de liaison à l'ADN, ce qui pourrait empêcher leur stabilisation et donc leur résolution structurale.

Le domaine de dimérisation de SopB (résidus 275-323) a été résolu de façon séparée. Cette région est composée de trois brins β antiparallèles et d'une hélice α en C-terminale. Comme pour ParB_{P1}, les hélices α de deux sous-unités de SopB interagissent ensemble, permettant ainsi la formation d'une structure étendue, composée de 6 feuillets β organisés de façon antiparallèle (Schumacher and Funnell, 2005; Schumacher, Mansoor, *et al.*, 2007) (**Figure 29b et d et Figure 30a et b**). La région de dimérisation de KorB est un peu différente des deux précédentes puisqu'elle présente cinq feuillets β antiparallèles avec un repliement proche de celui des domaines d'interaction protéine-protéine des protéines Src 3 (SH3) (Delbrück *et al.*, 2002) (**Figure 29e**). SpoOJ_{T.th}, pourtant un des homologues de ParB de P1, montre aussi des similitudes de séquence avec KorB au niveau de sa région C-terminale. Cependant, cette région étant absente du cristal de SpoOJ (1-122), il est impossible de savoir si le domaine de dimérisation de SpoOJ_{T.th} adopte un repliement à la manière de ParB de P1 ou de KorB (Leonard *et al.*, 2004). Malgré l'absence du domaine de dimérisation, un dimère de SpoOJ_{T.th} est tout de même observé dans la structure cristalline (Leonard *et al.*, 2004). Ce dimère est formé par des résidus de motif B principalement (**Figure 29c**). Il y a deux principales interfaces d'interaction au sein de ce dimère: l'une entre les résidus de $\alpha 1$ d'un monomère et les résidus de $\alpha 3$, $\beta 1$ et $\beta 2$ d'un autre; et la seconde entre les résidus de $\alpha 2$ d'un monomère et ceux des hélices $\alpha 4$, $\alpha 5$, $\alpha 8$ et $\alpha 10$ de l'autre monomère. Ceci est cohérent avec des études réalisées sur d'autres protéines ParB de type Ia, telles que ParB de P1, qui suggèrent que le motif B est impliqué dans l'interaction dimère-dimère de ParB_{P1} (Surtees, 2001).

Le domaine central (155-270) de SopB est constitué de sept hélices α comme pour ParB de P1. Cette région peut elle-même être sous divisée en deux régions fonctionnelles, une région contenant le motif HTH (hélice $\alpha 2$ et $\alpha 3$) et une région contenant les hélices $\alpha 6$ et $\alpha 7$ qui

dimérise avec les hélices homologues $\alpha 6$ et $\alpha 7$ d'un autre SopB (**Figure 29b**). En plus du domaine de dimérisation C-terminal de SopB, ces deux dernières hélices définissent un second domaine de dimérisation. Les auteurs suggèrent que celui-ci pourrait permettre à SopB de s'oligomériser avec d'autres SopB pour permettre de lier plusieurs molécules d'ADN (**Figure 29b**). En effet, dans la structure, il permet à un dimère de SopB d'établir des contacts spécifiques avec un ADN centromérique et des contacts non spécifiques avec deux autres duplex d'ADN via l'oligomérisation avec d'autres SopB. Ainsi, cette structure suggère un mécanisme par lequel SopB pourrait d'une part réaliser l'appariement des plasmides nouvellement répliqués mais également permettre à SopB de recouvrir l'ADN non spécifique. Contrairement à SopB, les régions centrales de KorB et ParB_{P1} ne participent pas à l'oligomérisation. L'hélice $\alpha 6$ de KorB est trop courte pour être impliquée dans la formation de dimères et celle de ParB_{P1} est tordue et son orientation ne permet pas de faire ce type de contact (Schumacher *et al.*, 2010).

Dans le cas de ParB de P1, les données issues de la structure ont permis d'identifier un autre domaine de fixation à l'ADN dans le domaine de dimérisation. La découverte de ce domaine de fixation permet de comprendre comment ParB peut interagir avec les multiples boîtes contenues dans les sites *parS* (**Figure 30b et c**). Chaque boîte est fixée par un module de fixation à l'ADN individuel. Les boîtes A sont reconnues par les motifs HTH, alors que les boîtes B sont fixées par le domaine C-terminal de dimérisation. Les deux modules de fixation à l'ADN n'interagissent pas entre eux et peuvent pivoter librement contactant ainsi de multiples combinaisons de boîtes A et B dans la boucle de *parS*. De plus, cette libre rotation ne permet pas seulement la liaison de ParB_{P1} aux différentes boîtes A et B, mais peut également permettre la liaison entre deux ADN différents, suggérant un rôle potentiel de ce mode d'interaction dans l'appariement de deux plasmides différents (Schumacher and Funnell, 2005) (**Figure 30a et b**).

Le dimère de SopB ne contient pas de deuxième motif de liaison à l'ADN au niveau de son domaine de dimérisation, comme ParB_{P1}. Cependant, la présence d'un deuxième motif de liaison au centromère n'est pas un cas particulier à ParB_{P1}, même si l'utilisation du domaine de dimérisation en est un. En effet, la structure du domaine de fixation de KorB sur son centromère est originale. Elle est constituée de huit hélices incluant la partie C-terminale du motif B, le motif HTH et une partie de la région reliant le motif HTH au domaine de dimérisation. L'hélice de reconnaissance du motif HTH de KorB n'apporte pas la spécificité d'interaction. Cependant,

les résidus Thr₂₁₁ et Arg₂₄₀ des hélices $\alpha 6$ et $\alpha 8$, qui sont à l'extérieur de l'hélice de reconnaissance du motif HTH, sont utilisées pour contacter des bases spécifiques du centromère. Pour ParB de *P. aeruginosa*, les résidus G198 et R219 situés dans les hélices $\alpha 6$ et $\alpha 8$ respectivement, sont également impliqués dans cette fonction (Kusiak *et al.*, 2011). A la différence de KorB, l'hélice de reconnaissance de ParB_{*P. ae*} est elle aussi impliquée dans la spécificité d'interaction ParB-centromère. Une étude récente réalisée sur le plasmide F d'*E. coli*, a également montré qu'un résidu arginine, R219, situé en dehors du motif HTH, est essentiel à l'interaction spécifique avec son centromère *sopC* (Sanchez *et al.*, 2013). Cette étude révèle une caractéristique conservée dans le règne bactérien : les protéines ParB contiennent un domaine étendu de liaison au centromère, composé de deux motifs séparés et essentiels. Ce point sera détaillé dans ce manuscrit dans le **Chapitre II des résultats**.

La région extrême N-terminale de SopB et de presque toutes les autres protéines adaptatrices de type **Ia**, est impliquée dans l'interaction avec leur protéine motrice (Hanai *et al.*, 1996; Surtees, 2001; Ravin *et al.*, 2003). KorB est une exception car il utilise son domaine de liaison à l'ADN pour se lier avec sa protéine motrice (Lukaszewicz *et al.* 2002). Par ailleurs, les différentes structures ont montré des domaines de multimérisation, dimère-dimère, assez différents de ceux qui étaient déjà connu. Par exemple, l'étude structurale de SopB, décrit ce domaine dans la partie centrale de la protéine alors que une étude antérieure le situe dans la région N-ter (Hanai *et al.*, 1996; Schumacher *et al.*, 2010).

L'absence de structure de la région N-terminale, empêche d'avoir une idée de l'architecture des complexes de partition de type **I**. A ce jour, on ne sait toujours pas si les complexes de type **Ia** forment des superhélices ou d'autres superstructures afin de stabiliser ou simplement d'interagir avec leurs ATPases et de permettre ainsi la séparation et la partition des plasmides comme le font les complexes de types **II** et **III**.

D. Le complexe de partition "étendu" pour le type **I**

Les études structurales faites sur les systèmes de partition de type **I** ne permettent pas d'envisager une architecture particulière de ces complexes. Cependant, l'idée d'un complexe d'ordre supérieur provient de trois constats et des études qui en résultent. Premièrement, les protéines ParB plasmidiques et chromosomiques de type **I** se fixent sur le centromère pour

former le complexe de partition. A partir de ce point de nucléation, les protéines ParB peuvent s'étaler sur plusieurs kpb de part et d'autre de la séquence *parS* afin de former un complexe nucléoprotéique étendu que nous appellerons "le complexe étendu". Cet étalement a été initialement mis en évidence par la capacité de SopB, d'éteindre l'expression de gènes situés à plusieurs Kb de distances dans un contexte de surexpression de la protéine. Deuxièmement, la liaison de la protéine adaptatrice ParB sur des séquences centromériques hautement répétées, est capable d'induire des changements topologiques dans la structure de l'ADN qui seront d'autant plus prononcés que les protéines ParB s'étalent sur la molécule d'ADN en dehors des sites *parS* (Biek and Shi, 1994; Hayes and Austin, 1994; Hoischen *et al.*, 2004). Troisièmement, les protéines ParB de type Ia forment toutes des foci qui colocalisent avec les plasmides ou l'origine des chromosomes dans le cas des ParB chromosomiques. Ces trois constats semblent indiquer que toutes les protéines SopB sont fortement localisées et recrutées autour du complexe de partition.

Les études sur la topologie des complexes nucléoprotéiques de partition de type I ainsi que sur l'étalement des ParB seront tour à tour détaillées dans cette section.

1. Etude topologique du complexe de partition

La première évidence de la formation d'un complexe de partition dit "étendu" provient de l'étude de l'interaction de SopB avec *sopC* du plasmide F. *In vivo*, la fixation de SopB sur un plasmide portant *sopC* (12 répétitions de 43 pb), induit un déficit d'environ douze supertours négatifs par rapport à ceux sans *sopC* (Lynch and Wang, 1994; Biek and Shi, 1994; Biek and Strings, 1995).

Il a été proposé que le déficit de supertours négatifs était dû à un enroulement de l'ADN contenant *sopC* autour de SopB, créant ainsi une zone de supertours positifs. Cependant, cette hypothèse a été réfutée par une étude faite au sein de notre équipe (Bouet and Lane, 2009). La fixation de SopB induit une courbure de 50° sur un ADN portant un motif *sopC* de 43 pb, cependant cette courbure n'est pas additive puisque la fixation de SopB sur deux motifs n'introduit qu'un angle de 70°. SopB qui ne se fixe que sur 10 des 12 sites *sopC* (Mori *et al.*, 1989; Pillet *et al.*, 2011), pourrait introduire au maximum un ou deux supertours négatifs, ce qui est insuffisant pour expliquer le déficit de douze supertours observés. L'autre possibilité est que SopB lui-même induise des supertours positifs en se fixant sur l'ADN. Or, il a été montré

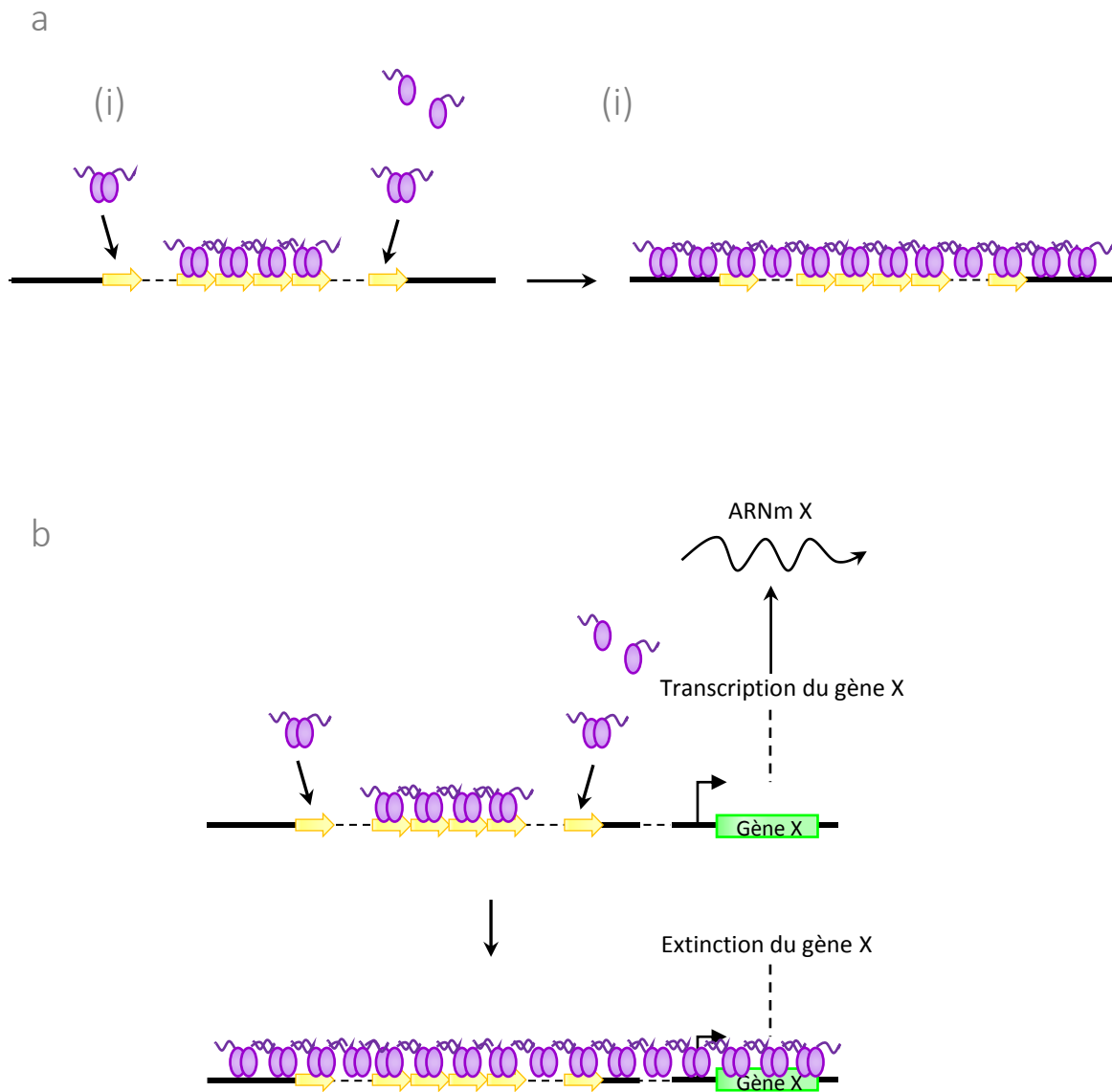


Figure 31 *Modèle de l'étalement des protéines ParB autour du site centromérique*

- a) Etallement des protéines ParB à partir de leur centromère. Les dimères de ParB (magenta) se fixent sur les sites centromériques (jaune) (i) puis s'évalent de manière non spécifique de part et d'autre de la séquence centromérique (ii). L'extrémité N terminale des protéines ParB, contenant le domaine présumé d'oligomérisation est flexible et est représenté par un bras en magenta.
- b) Modèle du "Silencing". Les dimères de ParB se fixent sur les séquences centromériques (en haut) puis s'évalent sur l'ADN conduisant à l'extinction des gènes avoisinants (gène X en vert) (en bas). Ces gènes sont normalement transcrits en absence de l'étalement.

que SopB ne pouvait pas introduire de supertours positifs, et ceci quel que soit l'ADN (ADN superenroulé ou relâché) et quelles que soient les conditions expérimentales (*in vitro* ou *in vivo*) (Bouet and Lane, 2009).

Cet effet sur la topologie de F existe aussi en présence d'une seule copie de 43 pb (déficit de 9 supertours négatifs). Ce déficit, comparable à celui observé en présence du site entier de *sopC* (12 supertours), suggère l'existence d'un complexe nucléoprotéique de grande taille qui pourrait être à l'origine d'une structuration à grande échelle du plasmide. Lynch et Wang ont donc proposé et montré que la fixation de SopB sur *sopC* puis son étalement sur les régions adjacentes pourrait expliquer le déficit de supertours observé. En effet, la fixation de SopB et son étalement semble protéger l'ADN de l'action de la gyrase (Lynch and Wang, 1995). Le complexe de partition formerait ainsi une structure suffisamment rigide pour résister à la transmission de changements topologiques depuis la région dépourvue de SopB vers la région concentrée en SopB.

Un excès de la protéine SopA est capable de réduire le déficit de supertours négatifs introduit par SopB. Ceci suggère qu'un excès de SopA interfère avec le complexe de partition "étendu" par interaction directe avec SopB et ceci, de façon ATP dépendante. En effet, le mutant SopA (K120Q), muté dans le site de liaison à l'ATP et supposé incapable d'interagir avec SopB, n'est plus en mesure de réduire le déficit de supertours négatifs (Lemonnier *et al.*, 2000). Ces résultats suggèrent que SopA serait impliquée dans la séparation des plasmides préalablement appariés au moment de la réaction de partition.

2. L'étalement de ParB sur l'ADN

L'étalement des protéines ParB de type Ia sur l'ADN non spécifique flanquant les sites centromériques est appelé "spreading" (**Figure 31a**). Ce dernier a été initialement mis en évidence par la capacité de SopB à éteindre l'expression d'un gène de résistance à un antibiotique placé au voisinage de *sopC* (Lynch and Wang, 1995; Kim and Wang, 1999). Les auteurs ont montré que SopB pouvait éteindre l'expression de gènes jusqu'à 10 Kb de distance. Ce phénomène est appelé "Silencing" (**Figure 31b**). Le "Silencing" a aussi été observé pour ParB de P1 (Rodionov *et al.*, 1999), KorB de RK2 (Jagura-Burdzy *et al.*, 1999; Bingle *et al.*, 2005), SopB de N15 (Grigoriev and Lobocka, 2001) ainsi que pour les homologues chromosomiques de *P. aeruginosa* (Bartosik *et al.*, 2004) et *B. cenocepacia* (Dubarry *et al.*, 2006). Des expériences

d'immunoprécipitation ont confirmé que le "Silencing" était la conséquence du "spreading" (Rodionov *et al.*, 1999). A cause de l'étalement de SpoB, l'ADN devient inaccessible aux autres protéines cellulaires. Dans le "Silencing", cela se traduit par la perte d'accessibilité de l'ARN polymérase aux régions promotrices (Williams *et al.*, 1993; Jagura-Burdzy *et al.*, 1999).

Dans toutes les conditions précédemment citées, le "silencing" a été détecté lorsque le ParB correspondant était surexprimé. Dans le cas de SpoOJ de *B. subtilis*, les quantités d'ARNm correspondants aux gènes situés autour de *parS* dans leur contexte chromosomique normal et en absence ou en présence d'un niveau physiologique de SpoOJ ont été comparés. Dans ce dernier contexte, aucun phénomène de silencing n'a pu être détecté. Cette dernière étude remet sérieusement en doute la pertinence biologique du silencing génique puisque ce phénomène ne semble pas se produire à des concentrations intracellulaires normales de ParB (Hao and Yarmolinsky, 2002; Breier and Grossman, 2007).

Par ailleurs, des expériences d'immunoprécipitation réalisées en condition physiologique, ont montré que les protéines ParB de P1, ParB de *S. coelicolor* et SpoOJ de *B. subtilis*, sont tout de même capables de s'étaler sur plusieurs Kb de part et d'autre de la séquence *parS* afin de former un complexe nucléoprotéique étendu (Rodionov *et al.*, 1999; Jakimowicz *et al.*, 2002; Rodionov and Yarmolinsky, 2004; Murray *et al.*, 2006; Breier and Grossman, 2007). Chez *B. subtilis*, ce complexe prend le nom de "domaine SpoOJ" (Murray *et al.*, 2006). L'étude du silencing a donc permis de mettre en avant la propriété des ParB à s'étaler sur l'ADN et a aussi servi de crible génétique pour détecter des mutants incapables de faire cet étalement (Lobocka and Yarmolinsky, 1996; Hanai *et al.*, 1996; Kusiak *et al.*, 2011). Avec les cribles génétiques, ces mutants ont été détectés sur toute la protéine, suggérant que la conformation globale de la protéine pourrait être importante pour réaliser le silencing. Il s'est avéré que certains d'entre eux étaient également déficients pour leur capacité de dimérisation ou de liaison à l'ADN, ce qui pourrait expliquer la perte du phénotype de silencing observé. L'étude par différentes techniques des protéines ParB et d'autres mutants, a permis de montrer que cet étalement pourrait se faire via la capacité de multimérisation des dimères de protéines ParB. Ce domaine de multimérisation semble être localisé au niveau du domaine N-terminal de la protéine (Hanai *et al.*, 1996; Kim and Wang, 1998; Kim and Wang, 1999; Surtees and Funnell, 1999; Bartosik *et al.*, 2004; Carmelo *et al.*, 2005). La protéine ParG de type **Ib** du plasmide TP228 peut également s'oligomériser lorsqu'elle est liée à l'ADN (Golovanov *et al.*, 2003).

Par conséquent, le "spreading" est une caractéristique commune à l'ensemble des ParB plasmidiques de Type I et chromosomiques mais son rôle exact dans le mécanisme de partition est encore flou. Une étude génétique sur la sélection de mutants déficients pour leur capacité de "Silencing" et de "spreading", montre qu'ils sont aussi déficients pour la partition, suggérant un rôle du "spreading" dans la partition (Lobocka and Yarmolinsky, 1996; Rodionov *et al.*, 1999). Chez *S. coelicolor*, la formation de ces complexes de partition "étendu" est nécessaire à l'organisation et à la ségrégation correcte des chromosomes dans les jeunes hyphes (Jakimowicz *et al.*, 2002; Jakimowicz *et al.*, 2005). Chez *B. subtilis*, le mutant SpoOJ93, défectueux pour sa fonction de spreading mais pas pour sa liaison au centromère, est incapable de réaliser correctement la ségrégation du chromosome et la sporulation, comme un mutant Δ SpoOJ (Piggot and Coote, 1976; Ireton *et al.*, 1994). Chez *P. aeruginosa*, des mutants équivalents, provoquent un défaut de ségrégation du chromosome, un ralentissement du taux de croissance, une mobilité réduite et de façon similaire aux mutants Δ ParB. Dans tous les cas testés, le spreading des ParB chromosomique semble important pour la viabilité cellulaire lors de certains stades de croissance.

L'importance du "spreading" a été remise en cause par une étude ultérieure qui consistait à limiter le "spreading" de ParB de P1 par l'introduction de barrières physiques de type "roadblock". Dans cette expérience, des sites de fixation des protéines RepA et GAL4 sont introduits de part et d'autre du site *parS* puis la stabilité du plasmide au fil des générations est mesurée (Rodionov and Yarmolinsky, 2004). Au vu du peu de perte obtenu après introduction de ces "roadblocks", il semble que le "spreading" ne soit pas essentiel pour la partition. Il est possible que l'espace laissé (222 pb) entre les deux "roadblocks" soit suffisant pour réaliser l'étalement nécessaire à la partition ou que les moyens de détection par PCR classiques, utilisés ici, ne soient pas assez sensibles pour conclure sur l'efficacité de ces barrières.

En revanche, le léger effet observé sur la partition laisse penser que la formation d'un complexe de partition "étendu" pourrait faciliter les étapes ultérieures du processus de partition, notamment l'appariement des plasmides et le recrutement de la protéine ParA aux systèmes de partition.

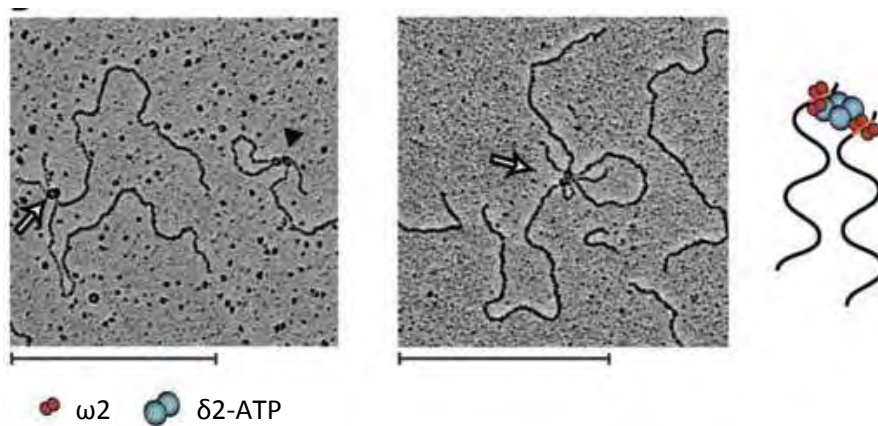


Figure 32 *Appariement intramoléculaire entre deux complexes de partition portés par le plasmide pSM19035*

Un ADN linéaire contenant *parS* (1 nM) est incubé avec 100 nM de $\delta 2$ (ParA), 60 nM de $\omega 2$ (ParB) et 1 mM d'ATP. Les ADNs appariés sont indiqués par les flèches vides. Les barres indiquent 500 nm. Adaptée de (Pratto *et al.*, 2008).

3. Rôle du complexe de partition "étendu"

a) L'appariement des plasmides

L'appariement des plasmides a été initialement suggéré car chaque copie du plasmide doit être correctement positionnée de sorte que chaque copie du plasmide soit dirigée dans chacune des cellules filles. Cette hypothèse est d'autant plus séduisante qu'elle rappelle l'image des chromosomes eucaryotes appariés lors de la métaphase pendant la mitose, suggérant un processus similaire pour les procaryotes. L'appariement des plasmides donne aussi une explication à l'incompatibilité. En effet, les plasmides qui utilisent des mécanismes similaires ou identiques de ségrégation ne peuvent coexister dans une même cellule car il s'exclue l'un l'autre (Austin and Nordström, 1990; Bouet, Nordström, *et al.*, 2007).

Le complexe de partition "étendu" constitué par l'étalement de ParB pourrait être responsable de l'étape d'appariement des plasmides. Quelques évidences viennent étayer cette hypothèse. En premier lieu, la surexpression de ParB de P1 et de SopB de F induit une perte supérieure à la perte aléatoire de plasmides ne portant pas de systèmes de partition (Kusukawa *et al.*, 1987; B E Funnell, 1988). L'une des explications est que l'excès des protéines ParB ou SopB pourrait empêcher la séparation des plasmides frères en induisant un appariement excessif. Une analyse a mis en évidence un effet recombinogène de la protéine SopB (Bouet *et al.*, 2006). En effet, en condition physiologique, SopB stimule la multimérisation par l'induction du système de recombinaison RecFOR. Cette recombinaison a pour substrat un ADN simple brin apparu lors d'une pause de la fourche de réplication induite par la présence de SopB. Ainsi, ce mécanisme de recombinaison pourrait favoriser un contact plus étroit et plus long entre les plasmides nouvellement répliqués. Ce résultat suggère indirectement l'existence d'un événement d'appariement des plasmides médié par SopB.

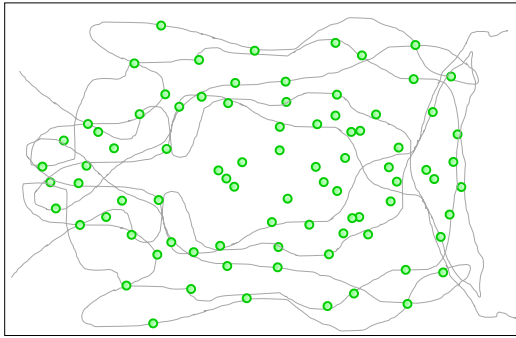
Aucune analyse de ME n'a permis, jusqu'à présent, de visualiser l'appariement de plasmides de type **Ia**. Pour les type **Ib**, l'appariement des plasmides pB171 au niveau de leur site centromérique a directement été observé par cette technique (Ringgaard *et al.*, 2007). Pour le plasmide pSM19035_{Ib}, une interaction intramoléculaire entre deux complexes de partition d'un même plasmide, a pu être mise en évidence (Pratto *et al.*, 2008) (**Figure 32**). Ceci suggère qu'un appariement entre deux complexes de partition portés par deux plasmides différents pourrait aussi être envisagé pour ce plasmide de type **Ib**. Il est à noter que dans le

cas des systèmes de partition de type **II**, l'appariement des plasmides R1 au niveau de leur site centromérique, *parC*, a pu également être observé en ME (**Figure 20**) (Jensen *et al.*, 1998). Ainsi, pour les plasmides de type **Ia** d'autres approches ont été envisagées afin de mettre en évidence un phénomène.

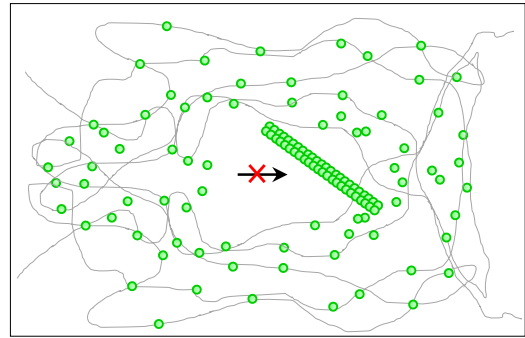
Pour le plasmide P1, une approche topologique *in vivo*, a permis de démontrer l'interaction entre deux sites *parS* portés par un même plasmide dimérique en présence d'une forte concentration de ParB. Cette interaction intramoléculaire entre deux sites *parS* montre que des événements d'appariement entre deux complexes de partition peuvent se produire (Edgar *et al.*, 2001). Les auteurs ont également montré que des mutants défectueux pour la capacité de "spreading" sont incapables de réaliser l'appariement des deux sites *parS*, suggérant ainsi que le complexe de partition "étendu" pourrait promouvoir l'appariement des plasmides.

Une approche de microscopie à fluorescence a aussi été envisagée pour voir l'appariement *in vivo*. Des observations faites sur les plasmides F, P1, pB171 et RK2 ont montré que chacun de ces plasmides est positionné en position médiane durant une partie importante du cycle cellulaire (Niki and Hiraga, 1997; Gordon *et al.*, 1997; Bignell *et al.*, 1999; Erdmann *et al.*, 1999; Pogliano *et al.*, 2001; Li and Austin, 2002; Ebersbach and Gerdes, 2004; Ebersbach *et al.*, 2006; Bouet, Nordström, *et al.*, 2007). Une fois répliqués, les plasmides sont rapidement séparés les uns des autres et repositionnés aux positions $\frac{1}{4}$ et $\frac{3}{4}$ de la cellule, ce qui les positionnent au centre des futures cellules filles. Les plasmides semblent ensuite se maintenir autour de ces positions données ce qui empêche leur ré-appariement éventuel (Derman *et al.*, 2008). Une étude réalisée sur le plasmide P1 montre que dans certaines cellules les foci effectuent des cycles de séparation/regroupement successifs, suggérant que l'appariement des copies du plasmide pourrait être un prérequis à la ségrégation et que plusieurs cycles de séparation pourraient avoir lieu par cycle cellulaire (Sengupta *et al.*, 2009). Bien qu'ils colocalisent les deux foci ne sont pas forcément dans le même plan ainsi cette étude ne permet pas de dire que les plasmides contenus dans ces foci se réappariaient. A ce jour, aucun événement d'appariement n'a réellement été observé *in vivo*, cependant l'avancée des techniques d'imagerie pourrait peut-être répondre à cette question.

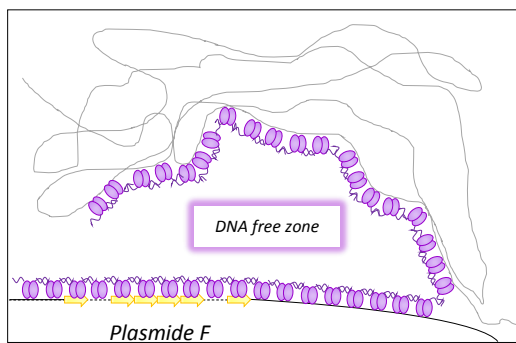
a



b



c



d

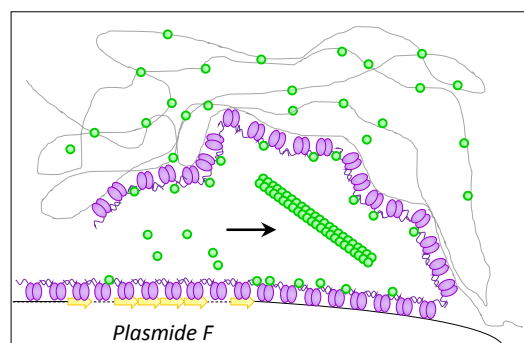


Figure 33 *Modèle de l'initiation de la polymérisation in vivo ou "DNA free zone"*

- a) SopA-ATP (vert) se lie au nucléoïde de manière réversible via son domaine d'interaction.
- b) Le nucléoïde inhibe la polymérisation de SopA.
- c) Quand le plasmide F est présent, SopB (magenta) se lie à *sopC*. A partir de ce point de nucléation, SopB s'étale sur le plasmide mais également sur l'ADN du nucléoïde pour créer une zone dépourvue d'ADN libre.
- d) Dans cette zone dépourvue d'ADN, l'inhibition causée par l'ADN libre est levée. SopA libre est recrutée par SopB. A cet endroit, SopA peut initier sa polymérisation et celle-ci conduira à la formation des polymères capables de réaliser la ségrégation du plasmide F.

b) L'interaction entre le complexe de partition et la protéine ParA

Les protéines ParB et l'assemblage du complexe de partition semblent être au cœur du mécanisme de partition. Des études ont montré, de manière directe ou indirecte, l'interaction du complexe de partition avec ParA via la protéine de liaison au centromère (Radnedge *et al.*, 1998; Bouet and Funnell, 1999; Lukaszewicz *et al.*, 2002; Figge *et al.*, 2003; Ravin *et al.*, 2003). En plus de faire le lien physique entre le complexe de partition et le moteur de la partition ParA, les protéines ParB, sont capables de réguler les activités de ParA. En effet, comme il a été vu précédemment, les protéines ParB stimulent la polymérisation de leur ATPase (Bouet *et al.* 2007) et l'activité ATPase de ParA augmente en présence d'un ou des deux partenaires du complexe de partition (ParB/*parS*) (cf. Introduction, Partie II.A.1 et II.A.1.a ; Figure 13). Ces données suggèrent que les différentes activités de l'ATPase ParA responsable du positionnement correct des plasmides, pourraient avoir une activité maximale là où la protéine ParB est concentrée, c'est à dire au niveau du complexe de partition et du complexe de partition "étendu".

Le complexe de partition "étendu" pourrait avoir un rôle direct dans le mécanisme de partition. Dans le cas du plasmide F, l'ATPase SopA est incapable de polymériser en présence d'ADN (Bouet, Ah-Seng, *et al.*, 2007). La présence importante d'ADN dans la cellule devrait inhiber la polymérisation de SopA. Cependant, il a été observé que SopB est capable de promouvoir la polymérisation de SopA *in vitro* (Bouet, Ah-Seng, *et al.*, 2007). De plus, la possibilité que SopB puisse recouvrir et masquer l'ADN grâce à ses capacités de liaison à l'ADN et d'étalement, pourrait permettre la polymérisation de SopA au niveau du complexe de partition. Au regard de ces données, notre équipe a proposé un modèle dans lequel SopB s'étalerait non seulement en *cis* sur l'ADN plasmidique ("*cis* spreading") mais également en *trans*, sur la surface du nucléoïde environnant le point de nucléation ("*trans* spreading") (Bouet, Ah-Seng, *et al.*, 2007) (Figure 33). Ce complexe (SopB-ADN) étendu pourrait former une zone appelée "DNA free zone" où l'ADN serait masqué par SopB. À cet endroit, SopA pourrait polymériser et amorcer le processus actif de ségrégation des plasmides. Ce modèle est en accord avec la structure du complexe SopB-sopC, montrant la capacité de SopB à réaliser du "*trans* spreading" en se liant à différentes molécules d'ADN.

Une récente étude sur le plasmide P1 a permis de mettre en évidence par un test de sédimentation et de diffusion de la lumière (Light Scattering), la présence d'un grand complexe ADN-protéine. La formation de ce complexe nécessite la présence de ParA, ParB, d'ADN centromérique et d'ATP (Havey *et al.*, 2012). L'assemblage et le désassemblage de ce complexe sont modulés par l'hydrolyse de l'ATP. Ce complexe est moins stable mais peut se former sur de l'ADN non spécifique et sa taille est directement proportionnelle à la longueur de l'ADN. Ces données sont compatibles avec le modèle de "réaction-diffusion" (**Figure 14**) dans lequel ParB, lié à l'ADN du plasmide P1, s'associe dynamiquement avec le nucléoïde bactérien via ParA (Vecchiarelli *et al.*, 2010). Puis, le nucléoïde sert de matrice à la partition des plasmides.

La capacité d'auto-assemblage des protéines ParB au niveau du centromère est une étape clé qui pourrait contrôler la partition en régulant les activités de ParA. A ce stade il apparaît nécessaire de préciser et de caractériser les différentes interactions impliquées dans la formation du complexe de partition et du complexe de partition "étendu". Le rôle exact du complexe étendu reste à ce jour très difficile à cerner, compte tenu du manque d'études moléculaires sur la nature et les caractéristiques de l'étalement de ParB sur l'ADN flanquant le centromère.

Les protéines ParB polymérisent-elles sur l'ADN à la manière d'un protéo-nucléofilament qui s'initierait au niveau du centromère? Dans ce cas là, la formation d'un protéo-nucléofilament pourrait créer une superstructure comme la super hélice du complexe ParR-*parC* de type **II** ou celle des complexes de type **III**, permettant leur interaction avec les polymères d'actine ou de tubuline, respectivement. L'étalement pourrait également ne pas correspondre à un protéo-nucléofilament. Dans ce cas-là, la fixation des ParB sur l'ADN flanquant les sites centromériques pourrait servir à compacter l'ADN. Ainsi, la fixation des ParB autour du centromère pourrait faciliter la compaction de la région origine dans le cas des ParB chromosomiques et celle du plasmide dans le cas des ParB plasmidiques. La compaction ou l'architecture adoptée pourrait aider le processus de ségrégation ou faciliter l'interaction avec un assemblage particulier des protéines ParA.

C'est dans ce contexte que s'inscrivent mes travaux de thèse. Quels sont les différents modes de fixation à l'ADN de ParB ? Avec qui et comment se font les différents types

d'interaction protéine-protéine pour conduire à la formation d'un complexe de partition " étendu ", substrat de la réaction de partition?

Résultats

Afin de mieux comprendre le mécanisme de la partition chez les bactéries, au sein de l'équipe nous développons comme modèle d'étude le système de partition de type Ia du plasmide F d'*E. coli*, appelé SopABC. Ce système Sop pour "Stability Of Plasmid" est composé de deux protéines, SopA et SopB, et comprend une séquence centromérique *sopC* située en aval de l'opéron *sopAB* (Ogura and Hiraga, 1983). Chacun de ces 3 éléments est essentiel à la partition du plasmide F (Ogura and Hiraga, 1983). SopA, l'ATPase du système, possède plusieurs activités. Par exemple, elle est capable de se fixer à l'ADN non spécifique, d'interagir avec SopB, de réguler la formation du complexe de partition "étendu" et d'osciller d'un pôle à l'autre du nucléoïde. Il a aussi été montré que SopA, ainsi que plusieurs homologues plasmidiques et chromosomiques dont ParF du plasmide TP228, ParA de pB171, ParA2 de *V. cholerae* et ParA de *C. crescentus*, polymérisent *in vitro*, en présence d'ATP (Barillà *et al.*, 2005; Ebersbach *et al.*, 2006; Bouet, Ah-Seng, *et al.*, 2007). L'activité ATPase de SopA est nécessaire à la partition. En effet, des mutations dans le domaine Walker box de SopA, entraînent un défaut de partition du plasmide F, comme pour les autres ATPase de type I étudiées (Davis *et al.*, 1996; Fung *et al.*, 2001; Libante *et al.*, 2001). Ces données permettent de supposer que l'ATPase SopA pourrait fournir l'énergie au déplacement des plasmides aux pôles opposés de la cellule. L'interconnexion et le rôle de ces différentes activités de SopA dans le processus de partition sont encore mal définis. SopB est la protéine de liaison à l'ADN qui reconnaît spécifiquement *sopC*. *sopC* est constitué de 12 répétitions d'une séquence de 43 paires de bases (pb) contenant un motif inversé-répété de 16 pb. Ce motif, reconnu spécifiquement par un dimère de SopB (Hanai *et al.*, 1996), permet l'assemblage d'une structure nucléoprotéique, appelée complexe de partition. A partir de ce point de nucléation, SopB va se propager sur l'ADN plasmidique de manière non spécifique (Lynch and Wang, 1995; Biek and Strings, 1995). En utilisant différents modes de fixation à l'ADN et en établissant des interactions protéine-protéine multiples, SopB est capable de s'organiser en complexe de partition dit "étendu". La formation du complexe de partition "étendu" est la première étape du mécanisme de partition et est essentielle au processus de ségrégation.

Au cours de ma thèse, j'ai initié plusieurs projets en parallèle, pour essayer de comprendre à la fois *in vivo* et *in vitro*, l'architecture adoptée par ce complexe et son rôle dans le mécanisme de polymérisation de la protéine SopA. Cette architecture n'est connue pour aucun des systèmes de partition ParABS de type I. Par des approches moléculaires et

biochimiques, je me suis intéressée aux différentes interactions impliquées dans la formation du complexe de partition, du complexe de partition "étendu" et comment il pouvait interagir et réguler les activités de SopA. Ce travail s'est déroulé en quatre parties.

Dans un premier temps, J'ai participé à des travaux sur la caractérisation des déterminants de spécificité de la séquence centromérique *sopC*, impliqués dans l'interaction avec la protéine ParB. Nous avons montré que SopB est uniquement capable de se lier spécifiquement à 10 des 12 répétitions du centromère *sopC*. Dans une répétition de 43 pb, seul le motif inversé-répété de 16 pb est nécessaire pour réaliser cette interaction. Les sept bases situées dans chacun des bras de l'inversé-répété sont requises pour l'interaction spécifique avec SopB et les bases A-C en tant que di-nucléotide purine-pyrimidine central, contribuent à l'efficacité de la liaison. En corrélation avec les études structurales réalisées sur SopB, ces travaux ont permis de proposer un modèle d'interaction entre l'HTH de SopB et le centromère *sopC*. Ces travaux sont présentés dans la première partie des résultats.

Je me suis ensuite intéressée aux déterminants de SopB impliqués dans l'interaction avec le motif de fixation centromérique de 16 pb. Nous avons identifié un nouveau motif en dehors du motif HTH et montré qu'une arginine de ce motif, l'arginine 219 pour SopB, est essentielle à l'interaction spécifique avec le centromère. Ces résultats ont révélé une caractéristique conservée dans l'ensemble du règne bactérien : les protéines ParB de type Ia possèdent un domaine étendu de liaison au centromère, composé de deux motifs séparés et essentiels. Ces travaux sont présentés dans la deuxième partie de mes résultats.

Ces deux premières études nous ont permis de comprendre les interactions impliquées dans l'assemblage du complexe de partition et les deux parties suivantes sont destinées à la compréhension de l'assemblage du complexe de partition "étendu". Par une combinaison d'approche moléculaire *in vivo* et *in vitro*, j'ai montré dans la troisième partie que l'architecture de ce complexe étendu en dehors des sites *sopC*, se faisait via deux types d'interaction : (i) des interactions protéine/protéine mais également (ii) ADN/protéine. Des données indirectes avaient montrées que SopB est capable de s'oligomériser et s'étendre sur l'ADN plasmidique de manière non spécifique. Dans cette étude nous avons montré de manière directe que SopB est capable de s'étaler sur plusieurs Kb de part et d'autre du centromère. Nous verrons que l'ensemble de ces résultats suggère que l'étalement de SopB ne résulte pas de la polymérisation de SopB sur l'ADN à la manière d'un protéo-nucléofilament qui s'initierait au

niveau du centromère. Ces travaux, nous ont permis de proposer un nouveau modèle dans lequel le complexe de partition "étendu" serait une structure concentrée, dynamique et résultant d'interaction stochastique entre SopB et l'ADN avoisinant *sopC*. De plus, des données préliminaires semblent indiquer que SopB serait capable de se fixer en "*trans*" sur l'ADN chromosomique, ce qui pourrait avoir un impact sur la conformation du complexe étendu mais également pourrait permettre la polymérisation de SopA *in vivo*.

La dernière partie des résultats concerne une étude préliminaire sur le rôle de l'étalement de SopB dans le mécanisme de partition et sur l'interaction de la protéine SopA avec le complexe de partition. Nous avons ciblé la région N-terminale de SopB, seule région dont la structure n'a pu être obtenue. Des déterminants de SopB impliqués dans l'interaction avec la protéine SopA ont pu être identifiés dans cette région. Nous avons montré que cette interaction permet de stimuler l'activité ATPase ainsi que la polymérisation de SopA. Nous verrons que la recherche du domaine d'oligomérisation de SopB s'est avérée plus délicate. Cependant, certains mutants semblent intéressants et pourraient servir pour des études ultérieures.

Chapitre I

Etude des déterminants nucléiques impliqués dans la spécificité d'interaction de la protéine SopB avec la séquence centromérique *sopC*

I. Etat de l'art

Le rôle de SopB dans le mécanisme de partition s'étend au-delà de sa capacité à servir de lien entre le centromère et SopA. SopB est capable de réguler toutes les activités de SopA. En effet, l'activité ATPase de SopA est très faible mais elle est stimulée par la présence d'ADN et de SopB (Watanabe *et al.*, 1992; Bouet, Ah-Seng, *et al.*, 2007). Une activité ATPase maximale est atteinte lorsque SopA interagit à la fois avec de l'ADN non spécifique et avec SopB liée sur le centromère *sopC* (Ah-Seng *et al.*, 2009). Ainsi l'activité d'hydrolyse est maximale au niveau du complexe de partition. La protéine SopB a également un double rôle dans la polymérisation de SopA : la protéine SopB stimule directement la polymérisation SopA et elle s'oppose également à l'inhibition de la polymérisation SopA par la présence d'ADN, en le masquant grâce à son activité de fixation à l'ADN non spécifique et probablement aussi par sa capacité à réaliser de l'étalement à partir du centromère *sopC* (Bouet, Ah-Seng, *et al.*, 2007). Ainsi, la formation des polymères de SopA pourrait être restreinte à l'endroit où SopB est fortement concentrée c'est-à-dire là où les complexes de partition "étendus" sont assemblés. Pour qu'il y ait formation d'un complexe de partition "étendu", il est absolument nécessaire que SopB soit spécifiquement assemblée sur les séquences centromériques *sopC*. Les déterminants de spécificité impliqués dans l'interaction SopB-*sopC* n'ont pas encore été déterminés et les bases qui sont nécessaires pour l'assemblage du complexe de partition de F ne sont pas connues.

sopC est constitué de 12 répétitions d'une séquence de 43 paires de bases contenant le motif inversé-répété de 16 pb sur lequel un dimère de SopB vient se fixer. Ce motif de liaison de SopB est fortement conservé au sein de la famille des SopB et ceci nous empêche de déduire quelles bases sont impliquées dans la spécificité de liaison SopB. Afin d'identifier au niveau des nucléotides les déterminants spécifiques nécessaires à l'assemblage du complexe de partition, nous avons développé une collaboration avec l'équipe de V. Leberre du laboratoire "Ingénierie

des Systèmes Biologiques et des Procédés" à Toulouse (INSA). Grâce à leur technique d'imagerie par résonance plasmonique de surface à haut débit (SPRi), nous avons pu faire une analyse exhaustive par variation des bases du motif de liaison (16 pb) de la protéine SopB. Une analyse des bases situées en dehors de ce motif et comprises dans la répétition de 43pb a également été réalisée. Cette approche nous a permis d'élaborer un modèle de fixation de SopB sur *sopC*. Une analyse des interactions ADN-protéine en retard de migration sur gel, nous a permis de préciser le modèle d'interaction SopB-*sopC* : (i) le motif inversé-répété de 16-bp de chaque répétition est suffisant pour assurer la spécificité d'interaction avec SopB bien que les bases des répétitions en tandem de 43 pb soient très conservées sur le centromère de F, (ii) toutes les bases du motif inversé-répété sont importantes à la fixation spécifique de SopB et (iii) les bases centrales non palindromique (Pu-Py) sont importantes.

L'association des résultats précédents et de la structure cristalline du complexe SopB-*sopC* a permis de préciser les contacts entre les bases de *sopC* et les acides aminés chargés positivement présents à l'intérieur (R190, R191 et R195) du motif hélice-tour-hélice (HTH) de SopB et de suggérer un contact possible avec un résidu présent à l'extérieur (R219) de ce motif (Schumacher *et al.* 2010). Un modèle plus précis de l'interaction SopB-*sopC* a pu être présenté dans l'étude qui suit.

La répétition du motif de liaison de SopB et l'espacement régulier entre chacun des motifs, pourraient être à l'origine d'une architecture particulière de complexe de partition. Nous nous sommes donc demandés si ces deux critères étaient conservés au sein de la famille des systèmes de partition de type Sop. Par une approche *in silico*, nous avons montré que l'espacement régulier est une caractéristique importante du centromère alors que le nombre de répétitions ne l'est pas. Par ailleurs, il a été montré qu'une seule répétition est suffisante pour permettre la formation d'un complexe nucléoprotéique et d'assurer le maintien au cours des générations du plasmide mini-F (Biek and Shi, 1994). La capacité de SopB de s'étaler sur les deux côtés de l'ADN centromérique pourrait compenser la variabilité dans le nombre de répétitions entre ces différents centromères et permettre la formation de complexes de partition d'une taille compatible avec une partition efficace.

II. Présentation de l'article :

Published online 7 June 2011

Nucleic Acids Research, 2011, Vol. 39, No. 17 7477–7486
doi:10.1093/nar/gkr457

Centromere binding specificity in assembly of the F plasmid partition complex

Flavien Pillet^{1,2,3}, Aurore Sanchez⁴, David Lane⁵, Véronique Anton Leberre^{1,2,3} and Jean-Yves Bouet^{4,5,*}

¹Université de Toulouse, INSA, UPS, INP, LISBP, 135 Avenue de Rangueil, F-31077 Toulouse, France, ²INRA, UMR792 Ingénierie des Systèmes Biologiques et des Procédés, F-31400 Toulouse, ³CNRS, UMR5504, F-31400 Toulouse, ⁴Université de Toulouse, Université Paul Sabatier, Laboratoire de Microbiologie et Génétique Moléculaires, F-31000 Toulouse and ⁵Centre National de Recherche Scientifique, LMGM, F-31000 Toulouse, France

Received March 3, 2011; Revised May 17, 2011; Accepted May 18, 2011

ABSTRACT

The segregation of plasmid F of *Escherichia coli* is highly reliable. The Sop partition locus, responsible for this stable maintenance, is composed of two genes, *sopA* and *sopB* and a centromere, *sopC*, consisting of 12 direct repeats of 43 bp. Each repeat carries a 16-bp inverted repeat motif to which SopB binds to form a nucleoprotein assembly called the partition complex. A database search for sequences closely related to *sopC* revealed unexpected features that appeared highly conserved. We have investigated the requirements for specific SopB–*sopC* interactions using a surface plasmon resonance imaging technique. We show that (i) only 10 repeats interact specifically with SopB, (ii) no base outside the 16-bp *sopC* sites is involved in binding specificity, whereas five bases present in each arm are required for interactions, and (iii) the A–C central bases contribute to binding efficiency by conforming to a need for a purine–pyrimidine dinucleotide. We have refined the SopB–*sopC* binding pattern by electro-mobility shift assay and found that all 16 bp are necessary for optimal SopB binding. These data and the model we propose, define the basis of the high binding specificity of F partition complex assembly, without which, dispersal of SopB over DNA would result in defective segregation.

INTRODUCTION

Faithful inheritance of chromosomal and extra-chromosomal DNA requires specialized molecular machines in eukaryotes and most prokaryotes (1,2). For low copy

number bacterial plasmids, loci ensuring active segregation to daughter cells before septum formation were identified over 30 years ago (3,4). Such partition loci are also present on most bacterial chromosomes (5). Virtually, all these loci specify three elements: a centromere, a centromere binding protein (CBP) and a NTPase. Interactions between these partners are essential for the partition process and of these, the specific recognition of centromeres by CBPs to assemble partition complexes is fundamental. CBPs act as the link between the centromere and NTPases and are thus functional equivalents of eukaryotic kinetochores. It is the partition complexes that are actively segregated by partition machinery.

The known partition loci can be classified into three types on the basis of the NTPase encoded: type I, Walker box (ParA); type II, actin-like (ParM); type III, tubulin (TubZ) (6). All types are found among plasmids, but type I is the majority and is the only type present on bacterial chromosomes (5). The most clearly understood system at the molecular level is type II, in which ATP enables the actin-like ATPase to form polymers that grow bi-directionally, moving the attached plasmids in opposite directions (7). Evidence is now accumulating that the type I segregation ATPases also polymerize in response to ATP binding and that this polymerization mediates DNA segregation (8–12). However, the molecular mechanism underlying this polymerization and how it contributes to DNA segregation by interaction with partition complexes have still to be unravelled.

Given the widespread distribution of type I partition systems and the central role of the partition complexes, it is important to define the determinants that allow the specificity of binding to centromeric sequences. Type I centromeres, generically termed *parS*, are diverse in size and organization (13). They are composed of small direct repeats, of inverted repeats (IR), or of more complex

*To whom correspondence should be addressed. Tel: (+33) 561 335 906; Fax: (+33) 561 335 886; Email: jean-yves.bouet@ibcg.biotoul.fr

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

© The Author(s) 2011. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

sequence arrangements as in the P1 centromere (14). In the case of IR sequences, they may be organized as tandem copies or present at different locations on the replicon. Chromosomal *parS* sites consist of a 16-bp IR motif as several copies, clustered close to replication origins (15) or spread throughout the origin domain as in *Bacillus subtilis* (16). Plasmid *parS* sites are typically restricted to a single locus. CBP binding to centromeres composed of the 16-bp IR motif occur generally through a single helix-turn-helix (HTH) domain, as observed for the F plasmid centromere (17). Nevertheless, the best characterized type I partition complex assembly is that of the P1 plasmid (18–20), consisting of an atypical centromere composed of two different classes of DNA repeat. The co-crystallization of a truncated version of P1 ParB with half of *parS* reveals that ParB binds as a dimer to the two binding motifs of *parS* through both HTH and dimerization domains (21), leading to more complex interactions than expected for centromeres composed of 16-bp IR motifs. In addition to specific binding, the CBPs also show significant non-specific DNA binding capacity that allows spreading to DNA adjacent to the centromere (22–24). Higher order organization of these extended partition complexes is still poorly understood for type I partition complexes.

The type I partition locus of the 100-kb conjugative plasmid F of *Escherichia coli* is composed of the ParA homolog, SopA, the centromere binding protein, SopB and the centromere, *sopC* (Figure 1). The *sopC* is organized as an array of 12 direct repeats of 43 bp (25), (see Figure 1 and Supplementary Figure S1), each containing a 16-bp IR motif. SopB is composed of several functional domains. A 3-D structure of the dimerization and HTH DNA binding domains, present in the C-terminal part of SopB, has recently been determined by X-ray analysis and a co-crystal structure of SopB (155–323)-*sopC* was also resolved at medium resolution (17). The N-terminal part of SopB is also multifunctional, but owing to its high flexibility, its structure is unknown (17). This part contains the SopA-interaction domain (26), the arginine-finger and the putative arginine-loop motifs that stimulate SopA ATPase activity (27) and possibly a SopB dimer dimer interaction region (28).

SopB binds to *sopC* with high affinity (29) through an HTH motif (17,26). Upon binding, it induces ~50° bend on the DNA (29), that if accumulated over the 12 *sopC*

sites could organize the F partition complex into a specific structure. The co-crystal structure of SopB (155–323) with *sopC* was however intriguing in two ways: (i) each of the HTH motifs of a SopB dimer contacts only one arm of a single *sopC* and (ii) *sopC* DNA is not bent; both are in contradiction with two previous studies (28,29). Like other ParBs, SopB not only assembles specifically on its centromere but also extends the complex by spreading along the DNA on either side of *sopC* (22); it might also spread *in trans* to different DNA molecules (17). *In vivo*, SopB bound to *sopC* and flanking DNA prevents negative supercoiling of the bound DNA (29,30), suggesting that the F partition complex shields DNA from the action of other proteins.

The role of SopB binding to DNA extends beyond that of a mere adaptor. The ATPase activity of SopA, as of ParAs generally, is very weak but is stimulated by DNA in the presence of SopB (10,31). Maximum activation is achieved when SopA interacts both with non-specific DNA and with SopB assembled on *sopC* (27). High levels of ATP hydrolysis by SopA thus occur close to assembled partition complexes. In addition, SopB has two roles on the formation of SopA filaments *in vitro*: it directly stimulates SopA polymerization and it counteracts the DNA inhibition of SopA polymerization by masking DNA in the vicinity of the centromere through the spreading/oligomerization activity (10). These roles of SopB are thought to restrict formation of SopA polymers to the place where partition complexes are assembled. Thus, SopB-mediated stimulation of SopA activities is effective where SopB molecules are concentrated, i.e. around the *sopC* nucleation point. Since SopB binds non-specific DNA, it is crucial that partition complexes are specifically assembled only on centromeric sequences. Specificity determinants involved in SopB-*sopC* interaction have not yet been determined and the bases that are required for the assembly of the F partition complex are thus not known.

The 16-bp SopB binding motif is strongly conserved and thus prevents us from deducing which bases are involved in SopB binding specificity. To identify at the nucleotide level the specific determinants needed to initiate F partition complex assembly, we first searched databases for sequences closely related to *sopC*. The sequences exhibited unexpected features. By using high-throughput surface plasmon resonance imaging (SPRi), we have characterized the DNA binding requirements for specific SopB-*sopC* interactions and have found that all bases of the *sopC* binding site are needed for maximal SopB binding strength.

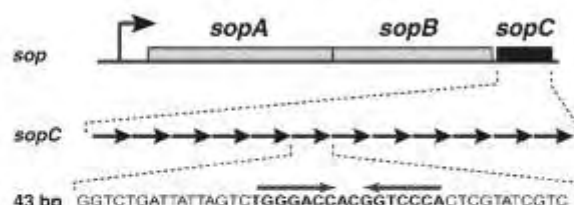


Figure 1. Organization of the F partition locus. The *sop* locus comprises two genes, *sopA* and *sopB* and a *cis*-acting centromeric site, *sopC*, composed of 12 direct repeats in tandem of 43-bp (represented by the repeated arrows). Each 43-bp unit contains a 16-bp SopB binding site. The inverted arrows indicate the SopB binding site. The bottom line shows the 43-bp consensus sequence (see Supplementary Figure S1 for the complete sequence alignment).

MATERIALS AND METHODS

Proteins and cell extracts

SopB was purified as previously described (10,27). Crude extracts containing SopB were produced from strain DLT812 (32) carrying the SopB expression vector pDAG170 (10), as follows. LB was inoculated by 100-fold dilution of an overnight culture of strain DLT812/pDAG170 and incubated at 37°C. At OD₆₀₀

~0.5, SopB production was induced by adding 0.1% arabinose. At 4 h after induction, cells were harvested by centrifugation, washed in TNE (50 mM Tris HCl, 50 mM NaCl, 1 mM EDTA), resuspended to an OD₆₀₀ ~200 in TNE supplemented with lysozyme (500 µg ml⁻¹) and NaCl (200 mM) and left on ice for 30 min. Viscosity was then reduced by sonication and cell-free extracts were obtained by centrifugation at 4°C for 20 min at 22 000g. Aliquots of crude extracts were then frozen rapidly in liquid nitrogen, allowing the same SopB extract to be used for all SPRI assays reported in this study. SopB was estimated by SDS-PAGE to constitute 10–15% of crude extract protein.

SPRI assays

SPRI analysis was performed using a SPRI-Plex system from GenOptics HORIBA-scientific (France). Synthetic 65-mer oligonucleotide probes, modified at their 5'-ends with a thiol function, were obtained from Sigma and spotted on SPRI-BiochipsTM covered with a 50 nm gold film specially developed for SPRI processing (prisms from GenOptics). Spots of thiol-modified oligonucleotides (25 µM) in spotting buffer (3× SSC, 450 mM NaCl and 45 mM sodium citrate, pH 7) were applied using a ChipWriterPro contact spotter (Biorad) with solid pin SSP015 from Arrayit Corporation. At least two spots of each oligonucleotide were deposited on prisms. Before reflectivity measurement, the best angle of incidence for each sensor-chip was chosen as described (33). SPRI experiments were done in R buffer (20 mM HEPES pH 7.4, 100 mM KCl, 50 µg ml⁻¹ salmon sperm DNA and 1 mM DTT) at a flow rate of 50 µl min⁻¹. Injections of target molecules (200 µl), either purified SopB or SopB-containing cell extracts diluted in R buffer, were carried out after saturation of the prism surface with Denhardt solution (Sigma). The reflectivity responses were obtained after subtraction of the averaged signal obtained on C1 probes. Reflectivity values were measured at the end of target injection (240 s), just before washing with R buffer. After each SopB interaction measurement, the surface was regenerated with NaOH 0.1 M. Data were analysed with Genoptics software SPRIAnalysis1.2.1.

Electro-mobility shift assays

Gel retardation assays were performed essentially as described (18), using 0.3 nM of annealed oligonucleotides (30 bp) 5'-³²P radiolabelled by polynucleotide kinase. Duplex oligonucleotides (E1, E2, E19, E23, E39, E40 and E46) were based on the *sopC* consensus: ATTAGT CtggaaccacgggtcccaCTCGTAT, with upper case letters being invariant and lower case letters representing the sequence and its variations (see highlighted nucleotides in Figure 6). Binding reactions were assembled on ice in a buffer containing 30 mM HEPES-KOH (pH 7.5), 100 mM KCl, 100 µg ml⁻¹ bovine serum albumin, 100 µg ml⁻¹ sonicated salmon sperm DNA, 10% glycerol and 1 mM dithiothreitol and incubated 15 min at 30°C. Products were resolved at 4°C by electrophoresis on pre-run 6% polyacrylamide gels in TBE (90 mM Tris borate, 1 mM EDTA) for 150 min at 160 V. Quantification of free

probes and shifted DNA complexes were quantified using MultiGauge software (Fugi). GraphPad Prism (V4.03) software was used for graphs, curve fitting and K_D calculation.

RESULTS

Centromeres related to *F sopC* are highly conserved

To define the bases required for specificity in the assembly of the partition complex on the *F* centromeric sequence, we used Blastn (NCBI) to search the nr nucleotide database for close homologs of *F sopC*, according to two criteria: (i) homology to the 16-bp SopB binding site, using several variations of the consensus binding site 5'-TGGGACCACGGTCCCA-3' and (ii) at least two tandem repetitions of the sequence. This search yielded 22 different *sopC*-related centromeres that conformed to the SopB binding site, i.e. none consisted of repeats of a minor variant. Sequence alignments are presented in Supplementary Figure S3 and data are summarized in Table 1. We also found several centromeric sequences with the *F* SopB binding site consensus on three bacteriophage genomes. These were not repeated in tandem but dispersed over the genome and for this reason they were not included in this analysis.

All sequences recovered belong to plasmids and none to chromosomes. They are present only in bacteria of the Enterobacteriaceae, indicating that Sop subfamily of centromeric sequences is not widespread amongst bacteria. The number of tandem repeats is highly variable, from 3 in pYVe227 (*Yersinia enterocolitica*) to 17 in pKN3 (*Klebsiella pneumoniae*). While most of the *sopC* centromeric sequences (16 over 22) are composed of 43-bp repeats, as for the *F* archetype, one centromere has 44-bp repeat units and five have 45-bp units. In total, we therefore had 205 SopB binding sites for analysis. Only two centromeres, those of pCVM and pLVPK, are composed solely of perfect 16-bp consensus binding sites; all others include variant SopB binding sites. Most of the variations occur on the first and last repeats, especially in smaller centromeres (Table 1). Only 4 of the 22 centromeres have consensus binding sites on the terminal tandem repeats.

The high divergence of the binding sites in the first and last repeats raises doubts as to their SopB binding capacity (also see below). Hence, we excluded these repeats from further analysis. Within the remaining 161 binding sites, only 34 base variations from the consensus were observed. Fourteen of these are variants of the central A-C dinucleotide that, notably, maintain the sequence Purine-Pyrimidine (Pu-Py) (Table 1). Of the 20 remaining variants, 8 are at the third position of the consensus site and the others at any position except the seventh.

This analysis reveals the central Pu-Py and the seventh base as probable important determinants of SopB binding. The strong conservation of the others, with the exception of position three, suggests they may be as well. Clearly, targeted mutagenesis was needed to further define the bases required for SopB binding.

Table 1. Summary of sequence conservation analysis of *sopC* centromere homologs

| Names (organism) | Repeats no. ^a (bp) | Consensus binding site ^a | Canonical first and last IR | Pu-Py central | Mutations at positions no. ^c 2,4,5,6,7 |
|------------------|-------------------------------|-------------------------------------|-----------------------------|---------------|---|
| pKPN3 (Kp) | 17 (43) | 14 | 1 | all | 1 |
| p1658 (Ec) | 16 (43) | 13 | 1 | all | 1 |
| pVM01 (Ec) | 13 (43) | 10 | 1 | all | 1 |
| pCVM (Se) | 12 (43) | 12 | 2 | all | 0 |
| pLVPK (Kp) | 12 (43) | 12 | 2 | all | 0 |
| pK2044 (Kp) | 12 (43) | 11 | 1 | all | 0 |
| pSC138 (Se) | 12 (43) | 11 | 2 | all | 0 |
| pAPEC-1 (Ec) | 12 (43) | 9 | 1 | all | 1 |
| pColBM (Ec) | 12 (43) | 9 | 1 | all | 1 |
| F (Ec) | 12 (43) | 8 | 0 | all | 2 |
| pYPTS01 (Ypt) | 11 (45) | 8 | 0 | all | 1 |
| pECOS88 (Ec) | 11 (43) | 8 | 1 | all | 1 |
| pU302L (St) | 10 (43) | 7 | 0 | all | 1 |
| pYV (Ypt) | 8 (45) | 6 | 0 | all | 0 |
| pKP187 (Kp) | 8 (44) | 7 | 2 | all | 0 |
| pVir68 (Ec) | 8 (43) | 6 | 0 | all | 0 |
| pCD1 (Yp) | 7 (45) | 5 | 0 | all | 0 |
| pYVa127 (Ye) | 6 (45) | 4 | 0 | all | 0 |
| pO157 (Ec) | 6 (43) | 4 | 0 | all | 0 |
| pCD (Yp) | 5 (45) | 3 | 0 | all | 0 |
| pETEC74 (Ec) | 4 (43) | 2 | 0 | all | 0 |
| pYVe227 (Ye) | 3 (43) | 0 | 0 | all | 0 |

Homologs of the *sopC* centromere were searched using Blastn (NCBI) and were found in bacteriophages closely related to *E. coli* (first column). When two or more identical *sopC* centromeres are present on different plasmids, only one is reported (see Supplementary Figure S3). The number of repeats present in each centromere of plasmids is indicated in the second column, with the length of the repeat in brackets. The numbers of perfect consensus IR SopB binding sites and of canonical first and last repeats are shown in third and fourth columns, respectively. IRs without a Pu-Py central position are numbered in the fifth column and the number of mutations compared with the consensus at positions important for SopB binding according to SPRi analysis (Numbers 2, 4, 5, 6 and 7; Figure 5) for all repeats of each centromere are shown in the last column. Ec: *E. coli*; Kp: *K. pneumoniae*; Se: *Salmonella enterica*; St: *Salmonella thyphimurium*; Ye: *Y. enterocolitica*; Yp: *Yersinia pestis*; Ypt: *Yersinia pseudotuberculosis*. pCVM, pColBM and pYVa127 are abbreviations of pCVM19633, pAPEC-ColBM and pYVa127/90, respectively.

^aNumber of repeats in each centromeric sequence with the size in nucleotides of each repeat indicated in bracket; bold values, highlight differences with the standard 43-bp repeats.

^bGGGACC Pu-Py GGTCACA.

^cFirst and last *sopC* repeats are excluded from this calculation; positions are numbered according to Figure 5.

SPRi setup for SopB-*sopC* specific interaction analysis

We performed SPRi experiments that allow efficient screening of a large number of sequence variations. Briefly, SPRi is based on SPR technology, with the SPR signal monitored by camera, allowing simultaneous analysis of several dozen interactions in real time (34,35). Oligonucleotides with a thiol group at their 5'-ends were directly coupled to a gold sensor-chip [see 'Materials and Methods' section]; (33)]. They were designed as palindromes to enable formation by self-annealing of 30-bp DNA duplexes with a single stranded hairpin at the apex (Figure 2A). Oligonucleotide C2 corresponds to the consensus SopB binding site while C1, taken as the reference (for subtracting background), has a random sequence with the same nucleotide composition as C2 (Figure 2B);

(see 'Materials and Methods' section). All the SPRi data presented in the following sections come from a single sensor-chip, preliminary settings having been established on three preceding ones.

When purified SopB was injected, a specific interaction was readily detected by plotting the differential response obtained with C2 and C1 probes (Figure 2C, left). The intensity of the response curve was proportional to SopB concentration. In parallel, we also performed SPRi assays with a crude extract of cells in which SopB had been overproduced (Figure 2C, right), to test the possibility of directly assaying a protein of interest without the need for purification. The specific SopB-*sopC* interaction was readily detectable and proportional to the amount of extract, indicating that SopB-containing crude extract could be further tested along with purified SopB to provide a complete comparison of these two different SopB preparations. In both cases, we verified that SopB does not bind specifically to single strand DNA at any concentration tested (Figure 2C; dotted lines).

We were not able to calculate accurate affinity constants for SopB-*sopC* interaction, because the sensitivity of the SPRi technique is not as high as the SPR using the Biacore system, with which we had been able to measure a K_D of 2.5 nM (27). However, the response intensity at a given SopB concentration, in the 150–400 nM range, is highly reproducible (see below). We therefore used the maximum signal response obtained in the differential response curve for probes carrying sequence variations to identify the determinants for SopB binding to *sopC*. All subsequent analyses were performed with both purified SopB and SopB-containing extract.

SopB does not bind the 1st and 12th repeats of *sopC* centromere

The first and last *sopC* repeats are the most divergent, both in *sopC* and the other F-related centromeres (Table 1 and Supplementary Figure S3). We assayed their binding properties using Probes C24 and C25, corresponding to 1st and 12th *sopC* repeats of F *sopC*, respectively (Figure 3). No interaction was detected with purified SopB or with SopB-containing extract. We also tested these binding sites in their natural context (Probes C26 and C27), in fragments equivalent to the 30-bp consensus probe, but again no interaction with SopB was detected. Hence, the 1st and 12th repeats do not bind SopB efficiently *in vitro*. This result agrees with a previous DNaseI footprinting experiment which showed that the first repeat was not protected by SopB (36).

The 16-bp IR is sufficient for SopB binding specificity

DNase I footprinting of *sopC* *in vitro* (36) and *in vivo* (29) indicated that SopB protected only the 16-bp IR sequences. However, the 43-bp tandem repeats of F are highly conserved, suggesting that sequences surrounding the IR might play some role in SopB binding. We tested this possibility with Probes C28 and C29, which carry three base changes immediately to the left and on both sides of the IR, respectively. SopB interaction was as efficient as with the consensus C2 in both cases (Figure 3).

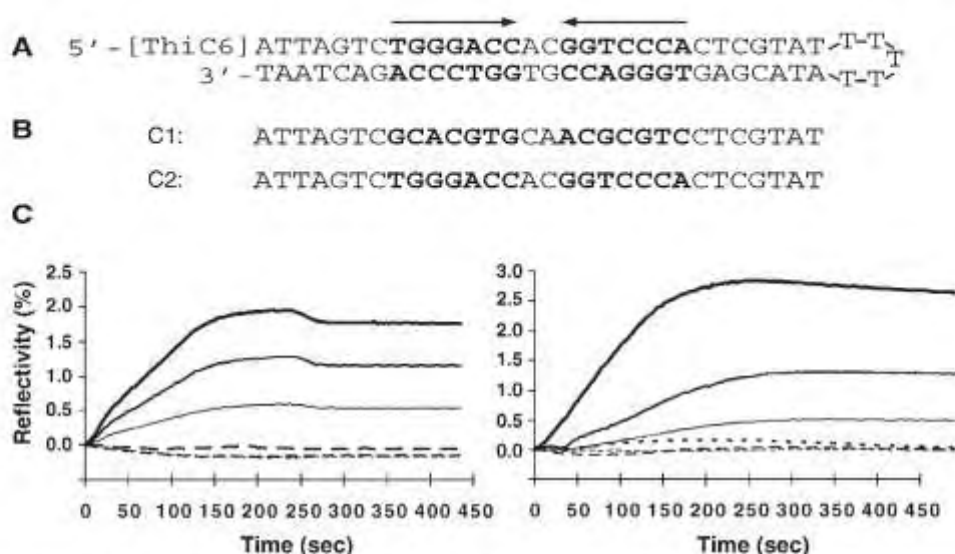


Figure 2. SPRi analysis of SopB binding to *sopC*. (A) Folding of the 5' thiol-labelled 65-mer single strand oligonucleotide; the sequence of oligonucleotide C2 is shown. (B) Sequences of the C1 (negative control) and C2 (consensus binding site) oligonucleotides. Note that the 16-bp consensus sequence present in C2 has been shuffled to generate C1. (C) SPRi dose response analysis. Results, expressed as the difference in reflectivity (%) between the signals obtained with the test and non-specific C1 probes and plotted as a function of time (seconds), were generated with purified SopB (left graph) or crude extract enriched in SopB (right graph). Buffer containing purified SopB at various concentrations (50, 100 and 150 nM; thin, medium and thick lines, respectively), or various dilutions of SopB extract (1/5000, 1/2000 and 1/1000; thin, medium and thick lines, respectively) were injected at time 0, followed by washing buffer (without SopB) after 200 s (see 'Materials and Methods' section). Full lines correspond to 65-mer C2 probes adopting a double strand folding and dotted lines to 30-mer C2 probes in single strand form.

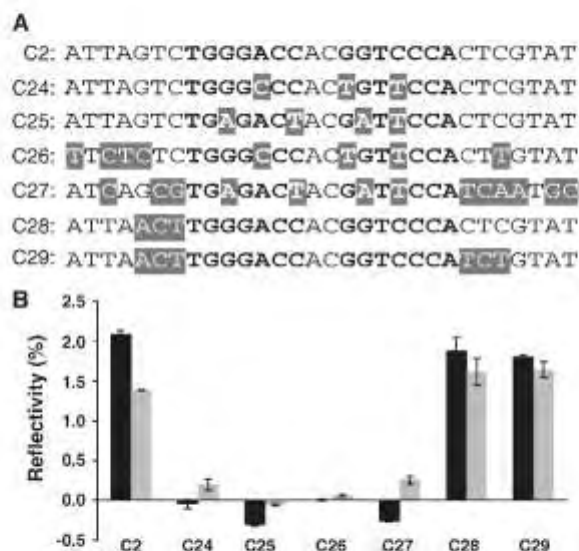


Figure 3. The 1st and 12th *sopC* repeats do not bind SopB. (A) Relevant sequences of analysed SPRi probes are shown below the consensus C2 probe. Left and right arms of the SopB binding site are in bold. Highlighted nucleotides indicate changes to the *sopC* consensus. (B) Histogram of SPRi signals for tested probes. Dark and grey bars represent the maximum of reflectivity (%) observed with purified SopB (250 nM) and with SopB-containing crude extract (1/500), respectively. Values are averages of duplicate readings on the same sensor-chip.

This result indicates that all site-specific binding determinants are present within the 16-bp IR sequence.

The central A-C is involved in SopB binding

The 16-bp IR includes a non-palindromic A-C central sequence. Only one of the twelve central dinucleotides of F *sopC* is different, A-T in the sixth IR (Supplementary Figure S1). To test whether these two central bases are important for the binding properties of SopB, as our *in silico* analysis had suggested (see above), we designed six probes with different central dinucleotides (Probes C3–C8; Figure 4). SopB interacted with Probes C3 and C7 about as strongly as with the consensus C2 probe, indicating that these modifications, C–T and A–G, do not affect binding. Notably, C3 corresponds to the naturally occurring A–T central sequence, suggesting that *in vivo*, this binding site is as efficient as the WT. All other probes tested (C4–6 and C8) showed a decrease in the SPRi signal of ~50%. This significant binding defect is found both with purified SopB and with SopB-containing cell extract. Probes C3 and C7 differ from the others in having a Pu–Py at the central position. These results, together with the observation that all the natural variations found in the central A–C conserve a Pu Py composition (see above), imply that an important requirement for SopB binding is a Pu–Py sequence at the centre of the binding site.

To confirm the importance of Pu–Py central bases for SopB binding, we used Probe C40, which consists of a

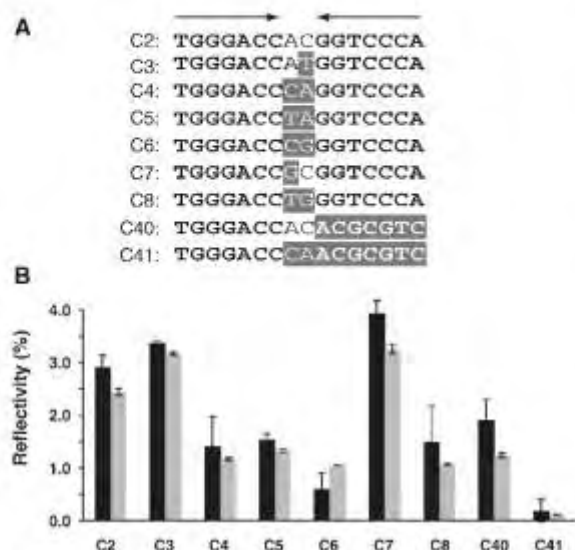


Figure 4. Non-palindromic A-C central bases are involved in SopB binding. (A) Sequences of analysed SPRi probes are shown below the consensus C2 probe as described in legend of Figure 3. Arrows on top represent the inverted-repeat motif. (B) Histogram of SPRi signals for tested probes. Dark and grey bars represent the maximum of reflectivity (%) observed with purified SopB (400 nM) or with SopB-containing crude extract (1/200), respectively. Values are averages of duplicate readings on the same sensor-chip.

half-binding site with the central A-C; the right arm being randomized as for Probe C1 (Figure 4). SPRi measurements indicated that C40 bound SopB less strongly (~50% of SPRi signal compared to C2). The equivalent probe with the central A-C changed to C-A (Py-Pu; Probe C41) showed no detectable interaction with either purified SopB or SopB-containing cell extract (Figure 4). Thus, combining removal of one arm with conversion of the central dyad to Py-Pu prevents stable complex formation, indicating that the changes are cumulative. These results clearly demonstrate that the Pu-Py central bases are necessary for fully efficient binding of SopB to the 16-bp IR motif.

Base pairs in *sopC* arms involved in SopB-*sopC* interaction specificity

To determine which bases in right and left arms of the 16-bp IR motif are involved in the specificity of binding, we made single base changes at each position or symmetrical base changes in both arms (positions are noted as 1-7; see Figure 5A). As shown above, removing a complete arm reduces to ~50% the SPRi signal obtained with SopB (Probe C40, Figure 4). Therefore, SPRi signals with probes carrying single changes are expected to be between 50% and 100% of that obtained with the consensus probe (C2). Indeed, we found this to be the case for all single base changes (Figure 5B).

Variations at Position 1 gave no significant reduction in SopB binding intensity when either single (C22) or double changes (C23) are made. Similar changes at Position 3

(C18 and C19) gave only minor reductions. Positions 1 and 3 thus appear to be unimportant for SopB binding specificity in this assay. By contrast, SopB interactions are significantly affected by single changes at Positions 2, 4, 5, 6 and 7 (Figure 5).

At two Positions (2 and 7) the reduction in signal depended on which arm carried the base change (Figure 5, compare C14-C145 and C20-C205). However, when changes were made on both arms for these two Positions (C15 and C21), SPRi signals were barely detectable, indicating that these positions are needed for SopB binding. At Position 4, symmetrical change (C11) resulted in an ~80% decrease in the SPRi intensity with both SopB and SopB-containing cell extract, indicating that this position is involved but not as strongly as Positions 2, 5-7. A discrepancy was observed with Probe C13, which retained ~20% of SPRi signal with purified SopB but was completely deficient when SopB-containing extract was used. This difference was not observed on a previous sensor-chip (no interaction was detected; data not shown), confirming that the double change on Position 6 abolished SopB interaction. These data indicate that the five Positions (2, 4, 5, 6 and 7) are important for SopB binding *in vitro*.

We also tested several combinations of multiple changes on one arm (Figure 5; Probes C42-C44). They all showed the same binding deficiency (~50% reduction in the SPRi signal) as observed when only one arm is present (Figure 4, Probe C40) and bound significantly better than a probe (C45) with variations in both arms.

It is interesting to note that amongst all the closely related *sopC* centromeres that have fewer than ten repeats, variations relative to the consensus are found exclusively on Positions 1 and 3 (Table 1 and Supplementary Figure S3; variations on first and last repeats are not considered). In addition, none of the three repeats present on the smallest centromere (pYVe227) have a consensus SopB binding site. However, variations on both the first and second repeats are restricted to Positions 1 and 3. These natural variations thus reinforce the results of the SPRi analysis in suggesting that Positions 2, 4, 5, 6 and 7 are the most important for SopB-*sopC* binding specificity.

Definition of the minimal SopB binding site

To confirm the binding specificity pattern observed in SPRi analysis, we performed electro-mobility shift assays (EMSA), which allow accurate measurement of SopB binding constants. DNA duplexes of 30 bp carrying the consensus or variant binding sites (see 'Materials and Methods' section) were incubated with increasing concentrations of SopB and subjected to EMSA analysis (Supplementary Figure S2) and binding curves were generated (Figure 6). About 50% of the *sopC* consensus probe (E2) was shifted at 3 nM of SopB (Figure 6, bottom), while no disappearance of the free probe carrying a shuffled sequence of the same nucleotide composition was observed (Probe E1; Supplementary Figure S2). Affinity constants (K_D) were estimated after fitting binding curves to a one-site binding hyperbola equation. SopB binds to the consensus 16-bp IR on the 30-bp probe with a K_D of

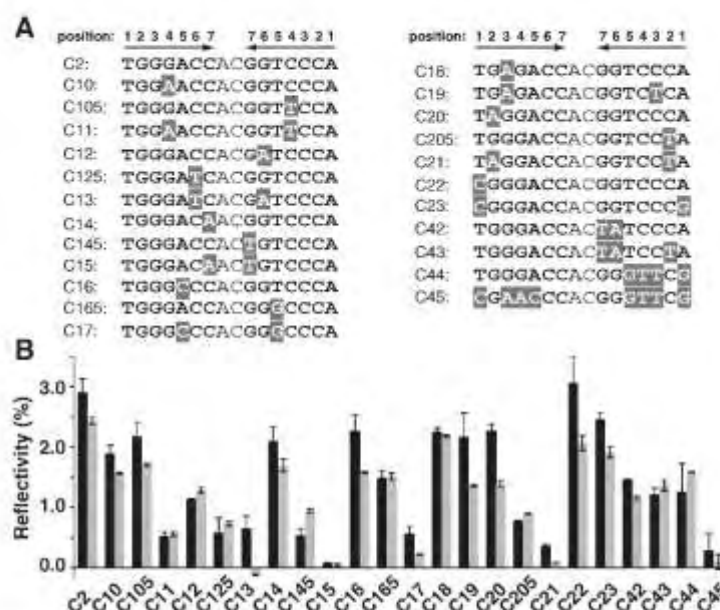


Figure 5. Five bases in each arm are required for efficient binding. (A) Sequences of analysed SPRi probes are shown below the consensus C2 probe as described in legend of Figure 3. Positions in both arms are numbered symmetrically from 1 to 7 due to their palindromic nature. (B) Histogram of SPRi signals for tested probes. Dark and grey bars represent the maximum of reflectivity (%) observed with purified SopB (400 nM) or with SopB-containing crude extract (1/200), respectively. Values are average of duplicate readings on the same sensor-chip.

2.6 nM, similar to values determined previously with longer DNA fragments containing a 43-bp repeat (27). No interaction was detected between SopB and Probe E45, equivalent to the SPRi Probe C45 (data not shown) or with Probe E46 carrying three variations on both arms at Positions 1, 3 and 4 (Supplementary Figure S2). When only two variations were present at Positions 1 and 3 on both arms (Probe E39), interaction with SopB was detected but binding affinity was decreased 15-fold ($K_D = 40$ nM; Figure 6). This loss of DNA binding efficiency was higher than expected from the SPRi data, which showed that individually these two variations did not significantly reduce SopB binding. We therefore measured binding affinities of *sopC* probes containing variations on both arms at Positions 1 (E23) or 3 (E19) and found SopB binding deficiencies of about 2.5- and 7-fold, respectively (Figure 6 and Supplementary Figure S2). These results are compatible with the SPRi measurements and indicate that binding deficiencies at Positions 1 and 3 are cumulative.

Model for SopB-*sopC* interaction

Our SPRi and EMSA assays measured only the specific binding of SopB to *sopC*. They were performed in the presence of non-specific DNA as competitor and in the case of SPRi analysis the binding curve was the differential response between signals obtained with tested and reference probes. We found that (i) five of seven bases in both arms of *sopC* are strongly involved in the specificity of SopB binding (Figure 5), (ii) bases at Positions 1 and 3 are required for optimal SopB binding (Figure 6) and (iii)

Pu-Py central bases contribute to efficient SopB-*sopC* interaction (Figure 4), revealing that all 16 bp in *sopC* site are involved in specific binding. A clear picture of the bases involved in specific binding of SopB to *sopC* can thus be generated (Figure 7).

Recently, the co-crystal structure of a SopB-*sopC* complex was resolved at about 3.5 Å (17), enabling prediction of likely contacts between *sopC* and positively charged amino acids present inside (R190, R191 and R195) and outside (R219) the HTH motif of SopB. Guanines at Positions 6 and 7, located close to amino acids R195 and R219, respectively, were found to be essential for binding in our SPRi analysis, strengthening the conclusion that interactions G6-R195 and G7-R219 are involved in SopB *sopC* binding specificity. However, the degree of resolution was too low to allow distinction between guanines G2 or G3 as the base in contact with the R190 residue. Our data enable this discrimination; mutation of G2 but not of G3 strongly reduces SopB binding, indicating that R190 interacts with the former. In the case of the K191 residue, a contact with either G4 or T5, or both was suggested by Schumacher *et al.* (17). Our results, indicating that variations on both positions have similar effects, are in agreement with K191 interacting with both G4 and T5 bases. These four positively charged residues of SopB are thus the main determinants of specificity for interaction with *sopC*. They are all conserved in the cognate SopB proteins corresponding to the *sopC* sequences listed in Supplementary Figure S3, as shown in Supplementary Figure S4. No contact was found in the SopB-*sopC* structure within the Pu-Py

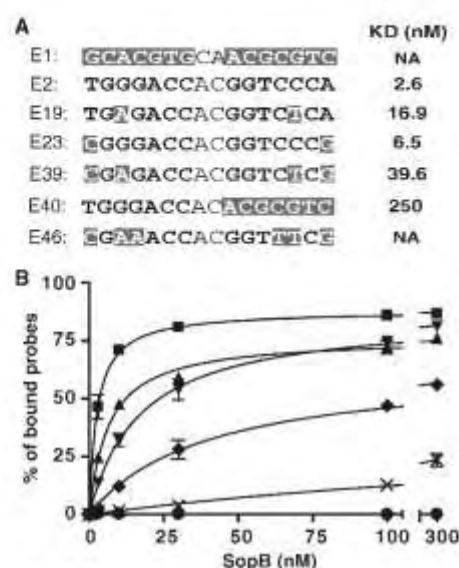


Figure 6. SopB *sopC* binding efficiency measured by EMSA. The 30-bp DNA duplexes carrying *sopC* or variant binding sites were incubated with increasing concentrations of SopB and subjected to EMSA (see Supplementary Figure S2). (A) Relevant sequences of analysed probes with highlighted nucleotides indicating changes to the *sopC* consensus (Probe E2). Binding constants (K_D) for variant *sopC* sites are calculated from the fitted binding curves according to the one site binding hyperbola equation $Y = B_{max} * X / (K_D + X)$. B_{max} is the maximum binding and K_D is the concentration of ligand required to reach half-maximal binding. (B) Quantification after EMSA of SopB binding to Probes E1, filled circles; E2, filled square; E19, filled inverted triangle; E23, filled triangle; E39, filled diamond; E40, cross. Each data point is the average of two independent experiments except for Probe E2 which was done four times. Binding curves are fitted to the above equation with $R^2 > 0.97$.

central position, suggesting that these two bases play a role in SopB-*sopC* interaction other than direct amino acid contact.

DISCUSSION

Assembly of partition complexes is the first step of the segregation process for bacterial replicons. As it is based on selective recognition of the centromere by the binding protein, knowledge of the determinants of this specificity is necessary for understanding partition at the molecular level. Using a large scale SPRi technique in parallel with EMSA, we have identified the DNA determinants of specificity in assembly of the partition complex on the F centromere. We have shown that the specificity determinants are restricted to the 16-bp IR and that all these bases are involved, to different extents, in the binding specificity to SopB (Figure 7). Our results point to an important constraint that prevents SopB binding at wrong positions. Such mis-assembly would lead to a segregation defect due to centromere incompatibility (37,38). In the course of this study on SopB-*sopC* interaction, we also demonstrated that crude extracts containing a protein of interest could

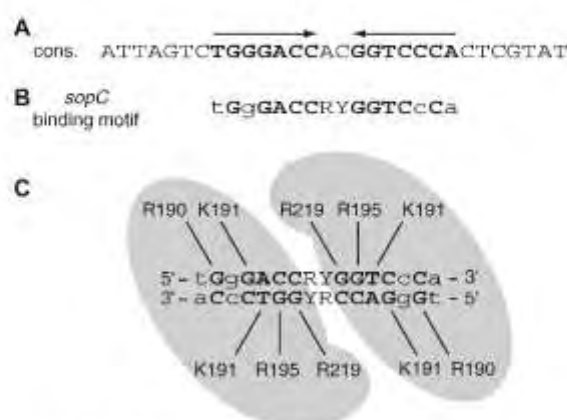


Figure 7. Refined *sopC* binding motif and model for SopB-*sopC* interactions. (A) The consensus sequence (cons.) of the *sopC* repeat. Arrows indicate the inverted repeat motif (bold nucleotides). (B) The SopB binding motif as defined by SPRi analysis. Essential nucleotides for SopB binding specificity are in bold upper case letters, important nucleotides are in upper case and nucleotides not strongly involved in specificity are in lower case. R and Y stand for purine and pyrimidine nucleotides, respectively. (C) Predicted interactions between amino-acids of SopB with *sopC* nucleotides. Grey zones represent the subunits of a SopB dimer. R190, R219, R195 and K191 correspond to arginines 190, 195, 219 and lysine 191. Straight lines represent suggested amino acid-nucleotide interactions.

be used in the high-throughput SPRi assay, as it proved to be as sensitive as highly purified protein in detecting DNA probes (Figures 2–5). Avoiding the need to purify proteins should be useful when the aim is to test binding of protein homologs or variants to numerous probes.

The number of *sopC* repeats in the Sop subfamily of partition system is highly variable, from 3 to 17 (Table 1), indicating that the actual number of repeats is not an important feature of the centromere. Indeed, it has been shown that a single repeat is sufficient to allow the formation of a highly structured nucleoprotein complex and to ensure stability of a mini-F (39,40). The ability of SopB dimers to spread on both sides of the centromeric DNA allows the formation of an extended partition complex (22,41). Spreading beyond the centromeric DNA may thus compensate for the variability in the number of *sopC* repeats by allowing the formation of partition complexes of a size compatible with efficient partition. It may be essential in the case of smaller centromeric sites found on some plasmids (Table 1). This could also be of particular importance in the case of chromosomal *parS* centromeric sites, which also consist of 16-bp IR but are scattered throughout in the origin region or clustered near the replication origin (15,16) rather than organized as tandem repeats as in *sopC* or other plasmid centromeric sites. Spreading of the centromere binding protein is indeed a property shared by centromere binding proteins encoded on chromosomes (15,23,42).

In vitro, the 1st and 12th *sopC* repeats do not bind specifically to SopB (Figure 3), due to variations on both arms at positions important for binding. This observation

is consistent with the bulk of the analysis (Figure 5) in showing that reducing binding on both arms completely prevents SopB binding. Only 4 of the 22 centromeres have consensus binding sites in their terminal repeats (Table 1) and centromeres with fewer than 11 repeats do not have a consensus binding site on either of these repeats, with one exception (pKP187). *In vivo*, these terminal repeats should not be bound directly by SopB. However, the interaction between SopB dimers that leads to spreading might enable the docking of SopB at the normal position on these defective sites. It will be of interest to determine whether the first and last *sopC* repeats of F are efficiently bound *in vivo* and whether the pattern of DNaseI protection is similar to that on the other repeats.

The SPRi analysis revealed the importance of the non-palindromic A-C central bases in SopB binding to *sopC* and showed that the important requirement is the presence of a Pu Py dinucleotide at these positions (Figure 4). Strikingly, all 205 SopB binding sites recovered from the database harbour a Pu-Py at the central position, even in the less conserved first and last repeats (Table 1 and Supplementary Figure S3), pointing to a key role of these two bases in SopB binding for all centromeres related to the Sop sub-family. However, in the recently published structure of SopB *sopC* complexes (17), no residue of SopB appears to be close enough to these central bases for direct interaction. These co-crystal SopB-*sopC* structures were obtained in conditions allowing SopB dimer to bind two different DNA molecules (17), which is in disagreement with an earlier study based on use of a chimeric SopB which suggested that the SopB dimer binds to a single *sopC* site (28). Our data support the latter view in which each of the HTH motifs of a SopB dimer contacts one arm of a single *sopC* site (Figure 7). Our arguments are based on the observation that several single mutations in one arm strongly destabilize the SopB-*sopC* interaction observed in SPRi (Figure 5), reducing the binding constant (measured by EMSA; Figure 6) by 100-fold. If only one arm was engaged in the primary interaction, we would not expect such a drastic decrease in binding efficiency. Moreover, when the A-C central bases were switched to Py-Pu in a configuration with only one arm, we observed complete loss of binding (Figure 3), indicating that SopB binding to only one arm of the site is not the actual binding mode.

We previously measured a bending angle of ~50  introduced upon SopB binding to a single *sopC* repeat (29). This bending, not observed in the SopB-*sopC* structure of Schumacher *et al.* (17), would be expected to involve the binding of both HTH motifs of a SopB dimer and might also be important in the specificity of the interaction. In a model in which both HTH motifs of a dimer bound to a single *sopC* site, SopB may thus interact directly with the central Pu-Py bases. Another likely possibility is that these bases may impose a structural constraint on *sopC* DNA. The bases flanking the central ones make a Py-Pu-Py Pu sequence arrangement in almost all sequences analysed (Supplementary Figure S3). Interestingly, it was observed that successive base planes in pyrimidine-purine steps open the angle between them (43,44). This could allow an optimal positioning of the bases present in the

major groove involved in SopB interactions as well as establishing the bend.

Centromeres of the Sop sub-family are composed of variable numbers of repeats organized in direct tandem. Most SopB binding sites are precisely spaced in a 43-bp motif of relatively well conserved sequences as for plasmid F (Supplementary Figure S3). We found some exceptions—one centromere with 44-bp repeats and five with a 45-bp repeat arrangement (Table 1). Such organization with a precise spacing argues for an important role of the sequence between the 16-bp IR sites. However, our data clearly show that no base outside the *sopC* site is involved in the primary interaction with SopB (Figure 3). This suggests that the role of the regular spacing is rather in the global organization of the partition complex. The 43-bp spacing corresponds almost exactly to four helical turns of the DNA. Each *sopC* binding site is thus present on the same side of the DNA allowing each SopB dimer to be positioned regularly on the same face. In addition, the ~50  bending induced by each SopB bound to *sopC* could organize the F partition complex into a continuous protein super-helical array, wrapping the DNA about its positive convex surface to form a solenoid-shaped structure reminiscent of that observed for the type II partition complex (6). We suggest that a widespread complex architecture formed by different classes of Par system may reflect some maximally efficient configuration for interaction with partition NTPases. In order to understand the global architecture of the Sop partition complex it will be of interest to investigate more carefully the role of this regular spacing in the partition complex assembly both *in vivo* and *in vitro*.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank J. Rech for technical assistance, Franck Pasta and Yoan Ah-Seng for discussions and Franck Pasta and Fran ois Cornet for critical reading of the manuscript.

FUNDING

l'Agence Nationale pour la Recherche (2010 BLAN 1316 01 to J.-Y.B., 2006 BLAN 0280 01 to D.L.); la R gion Midi-Pyr n es (07005583 to V.L.); GenOptics Horiba Scientific (to F.P.); MERNT fellowship (to A.S.). Funding for open access charge: Agence Nationale pour la Recherche N  2010 BLAN 1316 01.

Conflict of interest statement. None declared.

REFERENCES

- Hayes, F. and Barilla, D. (2006) The bacterial segrosome: a dynamic nucleoprotein machine for DNA trafficking and segregation. *Nat. Rev. Microbiol.*, **4**, 133–143.
- Cheeseman, I.M. and Desai, A. (2008) Molecular architecture of the kinetochore-microtubule interface. *Nat. Rev.*, **9**, 33–46.

3. Ogura, T. and Hiraga, S. (1983) Partition mechanism of F plasmid: two plasmid gene-encoded products and a *cis*-acting region are involved in partition. *Cell*, **32**, 351–360.
4. Austin, S. and Abeles, A. (1983) Partition of unit-copy miniplasmids to daughter cells. I. P1 and F miniplasmids contain discrete, interchangeable sequences sufficient to promote equipartition. *J. Mol. Biol.*, **169**, 353–372.
5. Gerdes, K., Møller-Jensen, J. and Bugge-Jensen, R. (2000) Plasmid and chromosome partitioning: surprises from phylogeny. *Mol. Microbiol.*, **37**, 455–466.
6. Schumacher, M.A., Glover, T.C., Brzoska, A.J., Jensen, S.O., Dunham, T.D., Skurray, R.A. and Firth, N. (2007) Segrosome structure revealed by a complex of ParR with centromere DNA. *Nature*, **450**, 1268–1271.
7. Garner, E.C., Campbell, C.S., Weibel, D.B. and Mullins, R.D. (2007) Reconstitution of DNA segregation driven by assembly of a prokaryotic actin homolog. *Science*, **315**, 1270–1274.
8. Barilla, D., Rosenberg, M.F., Nobbmann, U. and Hayes, F. (2005) Bacterial DNA segregation dynamics mediated by the polymerizing protein ParF. *EMBO J.*, **24**, 1453–1464.
9. Lim, G.E., Derman, A.I. and Pogliano, J. (2005) Bacterial DNA segregation by dynamic SopA polymers. *Proc. Natl Acad. Sci. USA*, **102**, 17658–17663.
10. Bouet, J.Y., Ah-Seng, Y., Benmeradi, N. and Lane, D. (2007) Polymerization of SopA partition ATPase: regulation by DNA binding and SopB. *Mol. Microbiol.*, **63**, 468–481.
11. Pratto, F., Cieck, A., Weihs, W.A., Lutz, R., Saenger, W. and Alonso, J.C. (2008) Streptococcus pyogenes pSM19035 requires dynamic assembly of ATP-bound ParA and ParB on *parS* DNA during plasmid segregation. *Nucleic Acids Res.*, **36**, 3676–3689.
12. Placín, J.L., Lee, S.F., Garner, E.C., Toro, E., Eckart, M., Comolli, L.R., Moerner, W.E. and Shapiro, L. (2010) A spindle-like apparatus guides bacterial chromosome segregation. *Nat. Cell Biol.*, **12**, 791–798.
13. Ebersbach, G. and Gerdes, K. (2005) Plasmid segregation mechanisms. *Annu. Rev. Genet.*, **39**, 453–479.
14. Funnell, B.E. and Gagnier, L. (1993) The P1 plasmid partition complex at *parS*: II. Analysis of ParB protein binding activity and specificity. *J. Biol. Chem.*, **268**, 3616–3624.
15. Bartosik, A.A. and Jagura-Burdzy, G. (2005) Bacterial chromosome segregation. *Acta Biochim. Pol.*, **52**, 1–34.
16. Lin, D.C.H. and Grossman, A.D. (1998) Identification and characterization of a bacterial chromosome partitioning site. *Cell*, **92**, 675–685.
17. Schumacher, M.A., Piro, K.M. and Xu, W. (2010) Insight into F plasmid DNA segregation revealed by structures of SopB and SopB-DNA complexes. *Nucleic Acids Res.*, **38**, 4514–4526.
18. Bouet, J.Y., Surtees, J.A. and Funnell, B.E. (2000) Stoichiometry of P1 plasmid partition complexes. *J. Biol. Chem.*, **275**, 8213–8219.
19. Schumacher, M.A., Mansoor, A. and Funnell, B.E. (2007) Structure of a four-way bridged ParB-DNA complex provides insight into P1 segrosome assembly. *J. Biol. Chem.*, **282**, 10456–10464.
20. Surtees, J.A. and Funnell, B.E. (2001) The DNA binding domains of P1 ParB and the architecture of the P1 plasmid partition complex. *J. Biol. Chem.*, **276**, 12385–12394.
21. Schumacher, M.A. and Funnell, B.E. (2005) Structures of ParB bound to DNA reveal mechanism of partition complex formation. *Nature*, **438**, 516–519.
22. Lynch, A.S. and Wang, J.C. (1995) SopB protein-mediated silencing of genes linked to the *sopC* locus of *Escherichia coli* F plasmid. *Proc. Natl Acad. Sci. USA*, **92**, 1896–1900.
23. Murray, H., Ferreira, H. and Errington, J. (2006) The bacterial chromosome segregation protein Spo0J spreads along DNA from *parS* nucleation sites. *Mol. Microbiol.*, **61**, 1352–1361.
24. Rodionov, O., Lobocka, M. and Yarmolinsky, M. (1999) Silencing of genes flanking the P1 plasmid centromere. *Science*, **283**, 546–549.
25. Hilsberg, M. and Eichenlaub, R. (1986) Twelve 43-base-pair repeats map in a *cis*-acting region essential for partition of plasmid mini-F. *J. Bacteriol.*, **165**, 1043–1045.
26. Ravin, N.V., Rech, J. and Lane, D. (2003) Mapping of functional domains in F plasmid partition proteins reveals a bipartite SopB-recognition domain in SopA. *J. Mol. Biol.*, **329**, 875–889.
27. Ah-Seng, Y., Lopez, F., Pasta, F., Lane, D. and Bouet, J.Y. (2009) Dual role of DNA in regulating ATP hydrolysis by the SopA partition protein. *J. Biol. Chem.*, **284**, 30067–30075.
28. Hanni, R., Liu, R.P., Benedetti, P., Caron, P.R., Lynch, A.S. and Wang, J.C. (1996) Molecular dissection of a protein SopB essential for *Escherichia coli* F plasmid partition. *J. Biol. Chem.*, **271**, 17469–17475.
29. Bouet, J.Y. and Lane, D. (2009) Molecular basis of the supercoil deficit induced by the mini-F plasmid partition complex. *J. Biol. Chem.*, **284**, 165–173.
30. Lynch, A.S. and Wang, J.C. (1994) Use of an inducible site-specific recombinase to probe the structure of protein-DNA complexes involved in F plasmid partition in *Escherichia coli*. *J. Mol. Biol.*, **236**, 679–684.
31. Watanabe, E., Wachi, M., Yamasaki, M. and Nagai, K. (1992) ATPase activity of SopA, a protein essential for active partitioning of F-plasmid. *Mol. Gen. Genet.*, **234**, 346–352.
32. Bouet, J.Y., Bouvier, M. and Lane, D. (2006) Concerted action of plasmid maintenance functions: partition complexes create a requirement for dimer resolution. *Mol. Microbiol.*, **62**, 1447–1459.
33. Pillet, F., Thibault, C., Bellon, S., Maillart, E., Trévisiol, E., Vieu, C., François, J.M. and Anton-Leberre, V. (2010) Simple surface chemistry to immobilize DNA probes that significantly increases sensitivity and spots density of surface plasmon resonance imaging based microarray systems. *Sensors and Actuators B*, **147**, 87–92.
34. Scarano, S., Scuffi, C., Mascini, M. and Minunni, M. (2010) Surface plasmon resonance imaging (SPR)-based sensing: a new approach in signal sampling and management. *Biosens. Bioelectron.*, **26**, 1380–1385.
35. Maillart, E., Brengel-Pesce, K., Capela, D., Roget, A., Livache, T., Canva, M., Levy, Y. and Soussi, T. (2004) Versatile analysis of multiple macromolecular interactions by SPR imaging: application to p53 and DNA interaction. *Oncogene*, **23**, 5543–5550.
36. Mori, H., Mori, Y., Ichinose, C., Niki, H., Ogura, T., Kato, A. and Hiraga, S. (1989) Purification and characterization of SopA and SopB proteins essential for F plasmid partitioning. *J. Biol. Chem.*, **264**, 15535–15541.
37. Bouet, J.Y., Rech, J., Egloff, S., Biek, D.P. and Lane, D. (2005) Probing plasmid partition with centromere-based incompatibility. *Mol. Microbiol.*, **55**, 511–525.
38. Bouet, J.Y., Nordstrom, K. and Lane, D. (2007) Plasmid partition and incompatibility - the focus shifts. *Mol. Microbiol.*, **65**, 1405–1414.
39. Biek, D.P. and Shi, J.P. (1994) A single 43-bp *sopC* repeat of plasmid mini-F is sufficient to allow assembly of a functional nucleoprotein partition complex. *Proc. Natl Acad. Sci. USA*, **91**, 8027–8031.
40. Lane, D., Rothenbuehler, R., Merrillat, A. and Aiken, C. (1987) Analysis of the F plasmid centromere. *Mol. Gen. Genet.*, **207**, 406–412.
41. Biek, D.P. and Strings, J. (1995) Partition functions of mini-F affect plasmid DNA topology in *Escherichia coli*. *J. Mol. Biol.*, **246**, 388–400.
42. Breier, A.M. and Grossman, A.D. (2007) Whole-genome analysis of the chromosome partitioning and sporulation protein Spo0J (ParB) reveals spreading and origin-distal sites on the *Bacillus subtilis* chromosome. *Mol. Microbiol.*, **64**, 703–718.
43. Dickerson, R.E. and Drew, H.R. (1981) Kinematic model for B-DNA. *Proc. Natl Acad. Sci. USA*, **78**, 7318–7322.
44. Zhurkin, V.B. (1983) Specific alignment of nucleosomes on DNA correlates with periodic distribution of purine-pyrimidine and pyrimidine-purine dimers. *FEBS Lett.*, **158**, 293–297.

Supplementary data

Figure S1. Consensus sequence of the 43-bp *sopC* repeats. Alignment of the *E. coli* plasmid F centromeric sequence (Mori *et al.*, 1986) every 43 bases allow the determination of the consensus sequence, presented at the bottom. Inverted arrows indicate SopB binding sites, present in each repeats, and composed of two 7-bp inverted repeat sequences (in bold), separated by 2 central bases. Highlighted bases correspond to differences from the consensus. Repeats are numbered from 1 to twelve on the left.


| Repeat # | |
|-------------|--|
| 1 | 1- G C C T G A A T A T T C T C T C T G G G C C A C G T C C A C T G T A T C G T C |
| 2 | 44- G G T C T G A T A T A G C T G G G A C C A C G G T C C C A C T C G T A T C G T C |
| 3 | 87- G G T C T G A T T A T T A G T C T G G G A C C A C G G T C C C A C T C G T A T C G T C |
| 4 | 130- G G T C T G A T T A T T A G T C T G G G A C C A C G G T C C C A C T C G T A T C G T C |
| 5 | 173- G G T C T G A T A T A G C T G G G A C C A C G G T C C C A C T C G T A T C G T C |
| 6 | 216- G G T C T G A T T A T T A G T C T G G G A C C A G G T C C C A C T C G T A T C G T C |
| 7 | 259- G G T C T G A T T A T T A G T C T G G G A C C A C G G T C C C A C T C G T A T C G T C |
| 8 | 302- G G T C T G A T T A T T A G T C T G G A C C A C G G T C C C A C T C G T A T C G T C |
| 9 | 345- G G T C T G A T T A T T A G T C T G G G A C C A C G G T C C C A C T C G T A T C G T C |
| 10 | 388- G G T C T G A T T A T T A G T C T G G G A C C A C G A T C C C A C T C G T S T G T C |
| 11 | 431- G G T C T G A T T A T C G T C T G G G A C C A C G G T C C C A C T G T A T G T C |
| 12 | 474- G A T C A G A C T A T A G C C T G A G A C T A C G A T C C A T C A A T G C C T G T |
| |  |
| consensus | GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC |

Figure S2. SopB-*sopC* binding efficiency measured by EMSA. Increased SopB concentrations (0, 3, 10, 30, 100 and 300 nM) were incubated with 0.3 nM of 32 P-labelled DNA fragments. Various *sopC* sequences were used as probes (E1, E2, E19, E23, E39, E40 and E46 are described in the main text; Fig. 6). All reaction mixtures contained unlabeled sonicated salmon sperm DNA ($100 \mu\text{g}.\text{ml}^{-1}$) as non-specific competitor DNA. Position of discrete SopB-*sopC* complexes (B1) and free DNA duplexes (*sopC*) are indicated on the left. * indicates a secondary complex formed when unpurified oligonucleotides, but not when HPLC purified oligonucleotides (bottom panel, E2), were used for duplexes formation.

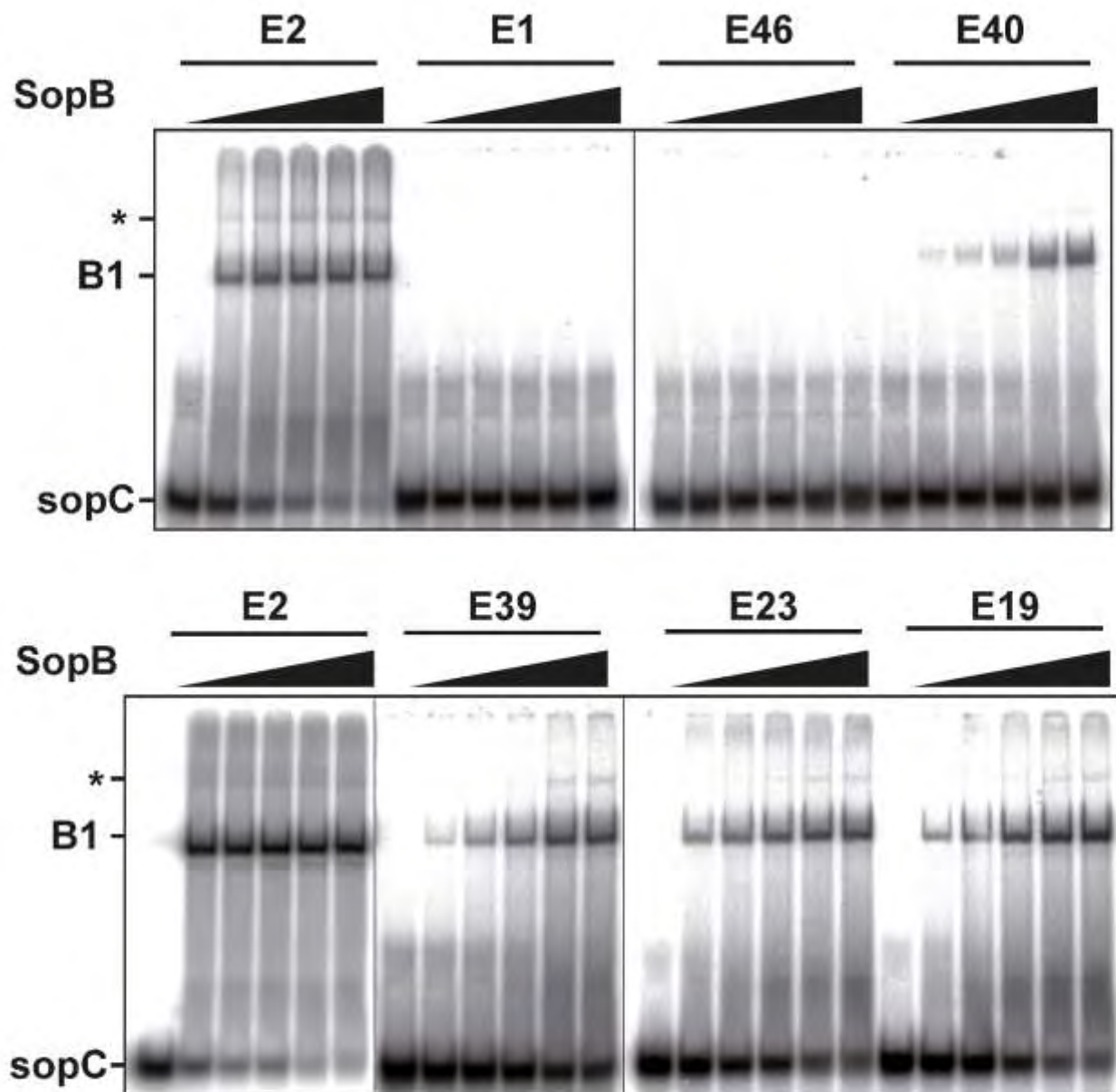


Figure S3: Alignments of centromeric sequences homologous to F *sopC*. Homologous sequences of the SopB (F) binding site consensus were identified in the nr database using Blastn (NCBI), with various combinations of query sequences. Sequences with at least 2 IR spaced by 43-, 44-, and 45- bp were retained for further analysis. Sequences were aligned to display repeats, and classified according to the number of repeats. Sequences in bold correspond to the SopB binding site. Nucleotides in blue correspond to the inverted repeat, while red ones are different from the consensus binding site.

- 17 repeats of 43 bp

Klebsiella pneumoniae plasmid pKPN3

>gi|150958208.gb|CP000648.1| *Klebsiella pneumoniae* subsp. *pneumoniae* MGH 78578 plasmid pKPN3

```

CCTTACTTTCATTTCCTTTCTCACAGGCCAGAAAACATAACTG
GCCTGAATATTCTCTCTGGGACCACGGTCCCACCCTGCATCGTC
GGTCAGGTTATCTGCCTGGGACCACGGTCTCACCCTGCCGCATC
GGTCAGGTTATCTGCCTGGGACCACGGTCCCACCCTGCCGCATC
GGTCAGGTTATCTGTCTGGGACCACGGTCCCACCCTGCCGCATC
TGTCAGGTAATCTGTCTGGGACCACGGTCCCACCCTGCCGCATC
TGTCAGGTAATCTGTCTGGGACCACGGTCCCACCCTGCCGCATC
GGTCAGGTTATCTGACTGGGACCACGGGCCCACCCTGCCGCATC
GGTCAGGTTATCTGCCTGGGACCACGGTCCCACCCTGCCGCATC
GGTCAGCTTATCTGTCTGGGACCACGGTCTCACCCTGCCGCATC
GGTCAGCTTATCTGTCTGGGACCACGGTCCCACCCTGCCGCATC
GGTCAGGTTATCTGTCTGGGACCACGGTCCCACCCTGCCGCATC
GGTCAGGTAATCTGCCTGGGACCACGGTCCCACCCTGTAATCGCC
GGTCAGGTTATCTGCCTGGGACCACGGTCCCACCCTGCCGCATC
GGTCAGGTTATCTGCCTGGGACCACGGTCCCACCCTGCCGCATC
GGTCAGCTTATCTGCCTGGGACCACGGTCCCACCCTGCCGCATC
TGTCAGCTTATCTGCCTGGGACCACGGTCCCACCCTGCATCGTC
GG CACATTATTAGTTAGGGATAACGGTCAATCAATGCCTGT
CACGAACAGGTATTGGCTTGTCGCCAGCAACCCGATGAGCGGA
    
```

(note that a gap was introduced in the 17th repeat for alignment)

- 16 repeats of 43 bp

Escherichia coli plasmid p1658/97

>gi|28629230.gb|AF550679.1| *Escherichia coli* plasmid p1658/97

```

TACTTAATTTCTTTAGCACAGGCTAGAAAACATAACTAACTG
GCCTGAATATTATATCTGGGACCACGGTCCCAAGTTGTATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCAACTCATATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCAACTCATATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCAAGTTGTATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCAACTCATATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCAACTCATATTGTC
GGTCTGATTATCGGTCTGGGACCACGGGCCCAACTCATATTGTT
AGTCTGCTAATCAGTCTGGGACCACGGTCCCAACTCGTATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCCCAACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCAACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCAACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCAACTCATATTGTT
GGTCTGATTATCGGTCTGGGACCACGGTCCCAACTCATATTGTT
AGTCTGATAATCAGTCTGGGACCACGGTCCCAACTCGTATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCTCACTTGTATTGTC
GATCTGATTATCAGCGTGAGACTACGATTCCATTAATGCCTGC
CAGTGAGTGTGTGGAACATACTTGGCCAGTGGCTTGGTCTGT
    
```

- 13 repeats of 43-bp

Escherichia coli plasmid pVM01

>gi|168830962|gb|EU330199.1| Escherichia coli plasmid pVM01.

```
TACTTAATTTTCCTTTAGCACAGGCTAGAAAACATAACTAACTG
GCCTGAATATTATATCTGGGACCACGGTCCCACTGTATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGATTATCGGTCGGGACCACGGTCCCACTCATATTGTT
AGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCGTATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTT
GGTCTGATTATCGGTCGGGACCACGGTCCCACTCATATTGTT
AGTCTGATAATCAGTCTGGGACCACGGTCCCACTCGTATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTTGTATTGTC
GATCTGATTATCAGCGTGAGACTACGATCCATTAAATGCCTGC
CAGTGAGTGTGTGGAACATATCTTGGCCAGTGGCTTGGTCTGT
```

(*S. enterica* pCVM29188_146 and *E. coli* pSMS35_130 have the same centromeric sequence as pVM01)

- 12 repeats of 43-bp

Salmonella enterica plasmid pCVM19633_110

>gi|194709275|gb|CP001125.1| Salmonella enterica subsp. enterica str. CVM19633 plasmid pCVM19633_110.

```
CCTTACTTTCATTTCCTTTCTCACAGGCCAGAAAACATAACTG
GCCTGAATATTCTCTCTGGGACCACGGTCCCACTGTATCGCC
GCTCAGGTTATCTCCCTGGGACCACGGTCCCACTGCGTCGTC
GGTCAGCTTATCTGCCGGGACCACGGTCCCACTGTATCGCC
GGTCAGGTTATCTGCCGGGACCACGGTCCCACTGCATCGCC
GGTCAGCTTATCTGCCGGGACCACGGTCCCACTGTATCGCC
GGTCAGGTTATCTGCCGGGACCACGGTCCCACTGCATCGTC
TGTCAGCTTATCTGCCGGGACCACGGTCCCACTGTATCGCC
GGTCAGGTTATCTGCCGGGACCACGGTCCCACTGCATCGTC
TGTCAGCTTATCTGCCGGGACCACGGTCCCACTGCTTCGTC
GGTCAGCTTATCTGCCGGGACCACGGTCCCACTGTATCGCC
GGTCAGGTTATCTGTCGGGACCACGGTGCCCACTGTATCGCC
GGTCAGGTTATCTGCCGGGACCACGGTCCCACTGCATCGTC
GGCACATTATTAGTTAGGGACAACGGTTCAATCAATGCCTGTC
```


***Klebsiella pneumoniae* plasmid pLVPK**

>gi|38016624|gb|AY378100.1| *Klebsiella pneumoniae* plasmid pLVPK.

```

CCTTTACTTCATTTCCTTTCTCACAGGCCAGAAAAATAACTG
GCCTGAATACTCTCTATGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGTTTTTCAGCCTGGGGCCACGGTCCCACCCTGCATCGTC
GTTTAGGCTTCCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GCTTAGGCTCCCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGCTCCCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGCTACCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGTTTTCCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGTTTC AGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGTTTTTCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGCTTCCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGTTTTTCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGTTTTTCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTACATTATTAGTTTGGGATCGCAGTTCATCAATGCCTGTC

```

(note that a gap was introduced in the 8th repeat for alignment)

***Klebsiella pneumoniae* NTUH-K2044 plasmid pK2044**

>gi|57158258|dbj|AP006726.1| *Klebsiella pneumoniae* NTUH-K2044 plasmid pK2044 DNA.

```

CCTTTACTTCATTTCCTTTCTCACAGGCCAGAAAAATAACTG
GCCTGAATACTCTCTATGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGCTTCCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GCTTAGGCTCCCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGCTCCCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGCTTCCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGTTTTCCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGTTTC AGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGTTTTTCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGCTTCCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGTTTTTCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTTAGGTTTTTCAGCCTGGGACCACGGTCCCACCCTGCATCGTC
GTTACATTATT AGTTTGGGATCGCAGTTCCAATCAATGCCTGT
CACGAACAGGTATTGGCATGGCGCCGGTGACCCGTTGAGCGGA

```

(note that gaps were introduced in the 7th and 12th repeats for alignment)

***Salmonella enterica* plasmid pSC138**

>gi|45758072|gb|AY509004.1| *Salmonella enterica* subsp. *enterica* serovar Choleraesuis str. SC-B67.

```

CCTTTACTTCATTTCCTTTCTCACAGGCCAGAAACATAACTG
GCCTGAATATTTCTCTTGGGACCACGGTCCCACCCTGCATCGCC
GCTCAGGTTATCTCCCTGGGACCACGGTCCCACCCTGCATCGTC
GGTCAGCTTATCTGCCTGGGACCACGGTCCCACCCTGCATCGCC
GGTCAGGTTATCTGCCTGGTGACCACGGTCCCACCCTGCATCGCC
GGTCAGCTTATCTGCCTGGGACCACGGTCCCACCCTGCATCGCC
GGTCAGGTTATCTGCCTGGGACCACGGTCCCACCCTGCATCGTC
TGTCAGCTTATCTGCCTGGGACCACGGTCCCACCCTGCATCGCC
GGTCAGGTTATCTGCCTGGGACCACGGTCCCACCCTGCATCGTC
TGTCAGCTTATCTGCCTGGGACCACGGTCCCACCCTGCTTCGTC
GGTCAGCTTATCTGCCTGGGACCACGGTCCCACCCTGCATCGCC
GGTCAGGTTATCTGTCTGGGACCACGGTGCCACCCTGCATCGCC
GGTCAGGTTATCTGCCTGGGACCACGGTCCCACCCTGCATCGTC
GGCATTATTAGTTAGGGACAACGGTTCATCAATGCCTGTC

```

(note that the 4th repeat contains 44 nucleotides for alignment)

gi221589217/gh-CP000836.1| Esch

[illegible]

TACTTAATTTCCCTTAGCACAGGCTAGAAAAACATAACTAAGT
GCGCTGAATATTATATCTGGGACCACGGTCCCACTTGATTGTG
GGTCTGATAATCAGTCTGGGACCACGGTCCCACTCATATTGT
GGTCTGATAATCAGTCTGGGACCACGGTCCCACTCATATTGT
GGTCTGATTATCGGTCTGGGACCACGGGCCCACTCATATTGT
AGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCGATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGT
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGT
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGT
AGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCGATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCTCACTTGATTGTT
GATCTGATTATCAGCGTGAGACACAGGATCCCACTTAATGCGCTGC
CAGTCAAGTGTCTGGAACATATCTGGCCACTGGCTTGCTGTG

gi|88770133|gb|DO381420.1| Escherichia coli

doi:10.1371/journal.pone.0172671.g002

TACTTAATTTCTCTTAGCACAGGCTAGAAAACATAACTAAGT
GCCTGAATATTATATCTGGGACCACGGTCCCACTTGTATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCACTCATATTGTT
GGTCTGATTATCGGTCGGGACCACGGGCCCACTCATATTGTT
AGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCGTATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTT
GGTCTGATTATCGGTCGGGACCACGGTCCCACTCATATTGTT
AGTCTGATAATCAGTCTGGGACCACGGTCCCACTCGTATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCTCCACTTGTATTGTC
GATCTGATTATCAGCGGTAGACTACGATTCATTAATGCGCTGC
CATCGAGTGTGTGAACATATCTTGGCCAGTGGCTTGGTCTG

1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 2678, 2679, 2680, 26

TCTTTACTTAATGTCCTTGTGTACAGGCCAGAAGCATAACTG
GCCTGAATATTCTCTCTGGGGCCCACTGTTCCA CTGTATCGTC
GGTCTGATTAATCAGACTGGGGACCACGGTCCC ACTCGTATCGT
GGTCTGATTATTAGTCTGGGGACCACGGTCCC ACTCGTATCGTC
GGTCTGATTATTAGTCTGGGGACCACGGTCCC ACTCGTATCGTC
GGTCTGATAATCAGACTGGGGACCACGGTCCC ACTCGTATCGTC
GGTCTGATTATTAGTCTGGGGACCACGGTCCC ACTCGTATCGTC
GGTCTGATTATTAGTCTGGGGACCACGGTCCC ACTCGTATCGTC
GGTCTGATTATTAGTCTGGAA ACCACGGTCCC ACTCGTATCGTC
GGTCTGATTATTAGTCTGGGGACCACGGTCCC ACTCGTATCGTC
GGTCTGATTATTAGTCTGGGGACCACGATCCC ACTCGTATTGTC
GGTCTGATTATCGGTCTGGGGACCACGGTCCC ACTTGTTATGTC
CATCAGACTATCACGCTGAGAC TACGATTCCATCAATSCCTGT
CBAAGSCAAGTATTGACATGTCTGCTTAACCTGTGTAACGGA

- 11 repeats of 45-bp

***Yersinia pseudotuberculosis* PBI/+ plasmid pYPTS01**

>gi|18670052|gb|CP001049.1| *Yersinia pseudotuberculosis* PBI/+ plasmid pYPTS01

```
CTAGTTTAAATCTCCCTAAAGCTATCAGTAAAAAATGCTCATCGA
CTTAATCGCTTTCATATCGAATCACATTCCACTAGGATGTTCAA
TGCCAGTGTAGTAATTGGGACCACGGTCCACCTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTGGATATCCAA
TGCCAACGTTAGTAATTGGGACCACGGTCCCACTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTGGATATCCAA
TGCCAACGTTAGTAATTGGGACCACGGTCCCACTGGATATCCAA
TGTCAACGTTGGTAATTGGGACCACGGTCCCACTGGATATCCAA
TGTCAACGTTGGTAATTGGGACCACGGTCCCACTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTGGATATCCAA
TGCCAACGTTAGTAATTGGGACCACGGTCCCACTGGATATCCAA
TGCTAACGTTGGTATGGGGCCACAGTTTCACTTAGGAAATGAG
TTGTAAAGAAGGAGATCTTCTACGAAGGATGAGCTAAACAATAG
```

- 11 repeats of 43 bp

***Escherichia coli* pECOS88**

>gi|218349681|emb|CU928146.1 *Escherichia coli* str. S88 plasmid pECOS88.

```
TACTTAATTTCTTTAGCACAGGCTAGAAAACATAACTAAGT
GCCTGAATATTATATCTGGGACCACGGTCCCACTTGTATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGATAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGATTATCGGTCGGGACCACGGGCCCACTCATATTGTT
AGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTC
GGTCTGCTAATCAGTCTGGGACCACGGTCCCACTCATATTGTT
GGTCTGATTATCGGTCGGGACCACGGTCCCACTCATATTGTT
AGTCTGATAATCAGTCTGGGACCACGGTCCCACTCATATTGTT
GGTCTGCTAATCAGTCTGGGACCACGGTCTCACTTGTATTGTC
GATCTGATTATCAGCGTGAGACTACGATTCCATTAATGCCTGC
CAGTGAGTGTGTGAACATATCTTGGCCAGTGGCTTGGTCTGT
```

- 10 repeats of 43 bp

***Salmonella typhimurium* plasmid pU302L**

>gi|37962716|gb|AY333434.1| *Salmonella typhimurium* plasmid pU302L.

```
TCTTTACTTAATGTCCTTTGTTACAGGCCAGAAAGCATAACTG
GCCTGAATATTCTCTCTGGGCCCACCTGTTCCACTTGTATCGTC
GGTCTGATAATCAGACTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GATCAGACTATCAGCGTGAGACTACGATTCCATCAATGCCTGT
CAAGGGCAAGTATTGACATGTCGTGTAACCTGTAGAAGGGAG
```

(*K. pneumoniae* pKF3-140, *E. coli* pIP1206 and *E. coli* pEK499 have same centromeric sequences as pU302L)

- 8 repeats of 45-bp

***Yersinia pseudotuberculosis* IP32953 pYV plasmid**

>gi|108860311|emb|BX936399.2| *Yersinia pseudotuberculosis* IP32953 pYV plasmid

```
TAGTTTAATCTCCCTAAAAGCTATCACTAAAAAATGCTCATCGA
CTTAATCGCTTTCAATTCGAATCACATTCCACTAGGATGTTCAA
TGCCAGTGTTAGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGTCACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTAGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTAGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCTAACGTTGGTGATGGGGCCACAGTTCAGCTTAGGAAATGAG
TTGTAAAGAAGGAGATCTTCTACGAAGGATGAGCTAAACAATAG
```

- 8 repeats of 44-bp

***Klebsiella pneumoniae* 342 plasmid pKP187**

>gi|206570325|gb|CP000965.1| *Klebsiella pneumoniae* 342 plasmid pKP187

```
CGGCCTCTCTAGTTTTTTTCTCAGGCCAGTTTTGATCAOCTGGC
CTTCATCTTGTGACGCCGGGACCACGGTCCCACTTCACCGTCA
GTACCGGAACCTCTGCCGAGACACGGTCCCACTTCACCGTCA
GCGTCGGAATTCTGCCGAGGACCACGGTCCCACTTCGCGCGTCA
GTGCCGGAATTCTGCCGAGGACCACGGTCCCACTTCGCGCGTCA
GCGCCGGAACCTCTGCCGAGGACCACGGTCCCACTTCGCGCGTCA
GTGCCGGAATTCTGCCGAGGACCACGGTCCCACTTCGCGCGTCA
GTACTGGAACCTCTTCCGAGGACCACGGTCCCACTTCGCGCGTCA
GCGCCGGAATGCTGCCGAGGACCACGGTCCCACTTCGCGCGTCA
GCGCCGGAATAGAGGTCATGTCTGTCATCGAATTGCTAATCGC
```

(note that a gap was introduced in the 7th repeat for alignment)

- 8 repeats of 43 bp

***Escherichia coli* plasmid pVir68**

>gi|253721152|gb|CP001162.1| *Escherichia coli* Vir68 plasmid pVir68.

```
TCTTTACTTAATGTCCTTTGTTACAGGCCAGAAAGCATAACTG
GCCTGAATATTCTCTCTGGGCCCACGTGTCCACTTGTATCGTC
GGTCTGATAATCAGACTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATCGGTCGGGACCACGGTCCCACTTGTATTATC
GATCAGACTATCAGCGTGAGACTACGATTCCATCAATGCCTGT
CAAGGGCAAGTATTGACATGTCGTGTAACCTGTAGAACGGAG
```

(*E. coli* pSFO157 has the same centromeric sequence as pVIR68)

- 7 repeats of 45-bp

***Yersinia pestis* Z176003 plasmid pCD1**

>gi|294352395|gb|CP001594.1| *Yersinia pestis* Z176003 plasmid pCD1

```
CTAGTTTAATCTCCCTAAAGCTATCACTAAAAAAATGCTCATCGA
CTTAATCGCTTTCAATTCGAATCACATTCCACTAGGATGTTCAA
TGCCAGTGTAGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGTC AACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTAGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCTAACGTTGGTGATTTGGGGCCACAGTTCAGCTTAGGAAATGAG
TTGTAAAGAAGGAGATCTTCTACGAAGGATGAGCTAAACAATAG
```

(*Yersinia enterocolitica* pYVe8081 has the same centromeric sequence as pCD1)

- 6 repeats of 45-bp

***Yersinia enterocolitica* strain A127/90 plasmid pYVa127/90**

>gi|28302111|gb|AY150843.2| *Yersinia enterocolitica* strain A127/90 plasmid pYVa127/90

```
TAGTTTAATCTCCCTAAAGCTATCACTAAAAAAATGCTCATCGA
CTTAATCGCTTTCAATTCGAATCACATTCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTGATTTGGGATCACAGTTCAGATTAGGAAATGAG
TTGTAAAGAAGGTGATCTTCTATGAAGGATGAGCTAAACAATAG
```

- 6 repeats of 43-bp

***Escherichia coli* plasmid pO157**

>gi|4589740|dbj|AB011549.2| *Escherichia coli* O157:H7 str. Sakai plasmid pO157 DNA.

```
TCTTTACTTAATGTCCCTTTGTTACAGGCCAGAAAGCATAACTG
GCCTGAATATTCTCTCTGGGGCCACGTTCCTTGTATCGTC
GGTCTGATAATCAGACTGGGACCACGGTCCCACTCGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTGTGTC
GGTCTGATTATCGGTCTGGGACCACGGTCCCACTCGTGTGTC
GGTCTGATTATCGGTCTGGGACCACGGTCCCACTTGTATTGTC
GATCAGACTATCAGCGTGAGACTACGATTCTATCAATGCCTGT
CAAGGGCAAGTATTGACATGTCGTCGTAACTGTAGAACGGAG
```

- 5 repeats of 45-bp

Yersinia pestis Pestoides F plasmid CD

>gi|145212871|gb|CP000669.1| Yersinia pestis Pestoides F plasmid CD

```
CTAGTTTAATCTCCCTAAAGCTATCACTAAAAAATGCTCATCGA
CTTAATCGCTTTCAATTCGAATCACATTCCACTAGGATGTTCAA
TGCCAGTGTTAGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCCAACGTTAGTAATTGGGACCACGGTCCCACTTGGATATCCAA
TGCTAACGTTGGTGATTGGGGCCACAGTTCAGCTTAGGAAATGAG
TTGTAAAGAAGGAGATCTTCTACGAAGGATGAGCTAAAACAATAG
```

- 4 repeats of 43-bp

Escherichia coli plasmid pETEC_74

>gi|157076671|gb|CP000799.1| Escherichia coli E24377A plasmid pETEC_74.

```
TCTTTACTTAATGTCCTTTGTTACAGGCCAGAAAGCATAACTG
GCCGTAATATTCTCTCTGGGGCCCACTGTTCCACTTGTATCGTC
GGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTGTGTC
GGTCTGATTATCGGTCGGGACCACGGTCCCACTTGTATTGTC
GATCAGACTATCAGCGTGAGACTACGATTCCATCAATGCCTGT
CAAGGGCAAGTATTGACATGTCGTCGTAACCTGTAGAACGGAG
```

(*E. coli* p557 has the same centromeric sequence as pETEC_74)

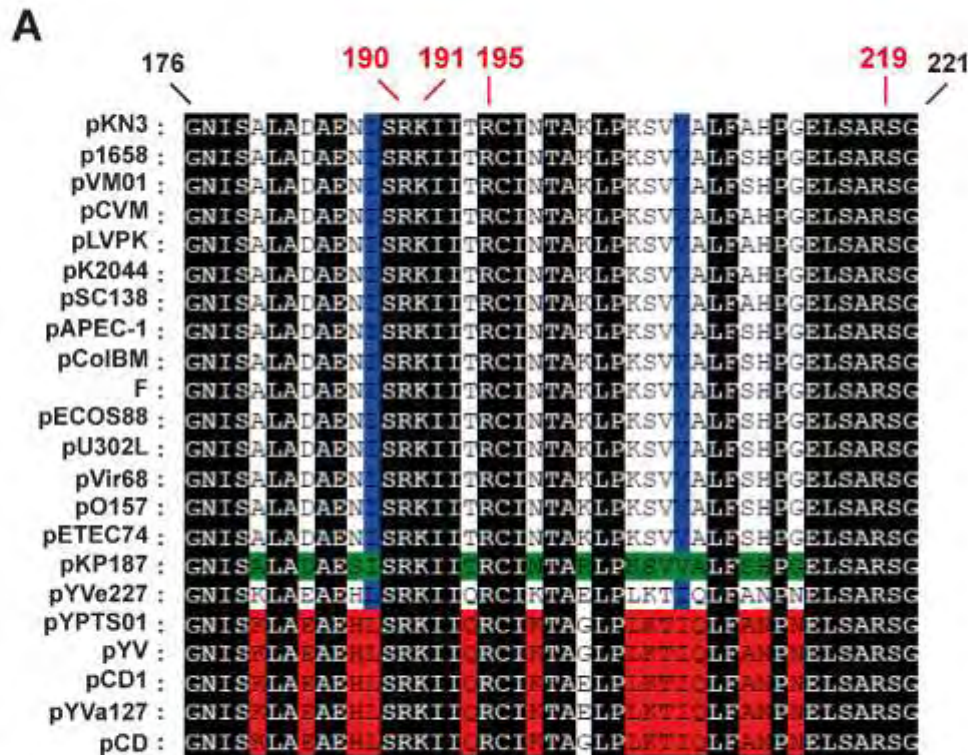
- 3 repeats of 43-bp

Yersinia enterocolitica plasmid pYVe227

>gi|4324323|gb|AF102990.1| Yersinia enterocolitica plasmid pYVe227.

```
TAATCGCTTTCAATTCGAATCACATTTCCACTTGGATATCCAA
TGCCAACGTTGGTAATTGGGACCACGGTCACCTGGATATCCAA
TACCAGCGTTAGTAATTGGGACCACGGTCACCTGGATATCCAA
TGCCAACGTTGGTGATTGGGACCACAGTTCAGCTTAGGAAATC
AATTGTAAAGAAGGACATCTTCAATGAAGGATGAGCTABAATA
```


Figure S4. Sequence conservation of the SopB specificity domain involved in SopB-*sopC* interaction. **A.** Alignment of SopB sequences, corresponding to cognate *sopC* sequences listed in Fig. S3, from residues 176 to 221. Residues conserved amongst all SopB members are highlighted in black. Residues conserved in subgroups belonging to 43-, 44- and 45-bp repeats carrying centromeres are highlighted in blue, green and red, respectively. Amino acid residues are numbered according to the F SopB sequence; R190, K191, R195 and R219 responsible for the specific binding to *sopC* 16-bp IR are indicated above in red. **B.** Weblogo of HTH motifs from all SopB sequences listed in A. SopB amino-acid sequences from residues 176 to 200 were aligned with ClustalW software and computed on the weblogo.berkeley.edu website to generate the sequence logo.



Chapitre II

Etude des déterminants des protéines ParB impliqués dans la spécificité d'interaction avec leur centromère : Caractérisation d'un second motif de liaison au centromère, essentiel pour le maintien du génome bactérien

I. Etat de l'Art

La protéine SopB comme toutes les CBPs de type **Ia** possède un motif Hélice-tour-hélice (HTH) de liaison au centromère. Ce motif est compris entre les résidus 178 et 198 de SopB. Le mutant SopB-A₁₈₃D-K₁₉₁A-R₁₉₅A a perdu sa capacité de fixation à l'ADN (Ah-Seng *et al.*, 2009). La mutation A₁₈₃D situé dans la première hélice de l'HTH, conduit à la perte de la fixation à l'ADN. Les mutations K₁₉₁A et R₁₉₅A situées dans l'hélice de reconnaissance, entraînent la perte de reconnaissance du centromère. L'association de deux récentes études, l'une sur la structure cristalline du complexe SopB-sopC et l'autre sur l'analyse des déterminants centromériques impliqués dans l'interaction SopB-sopC, a permis de préciser les contacts entre sopC et les acides aminés chargés positivement présents à l'intérieur (R190, R191 et R195) du motif HTH (Schumacher *et al.*, 2010; Pillet *et al.*, 2011) (**cf. Résultats, Chapitre I**). La structure a également prédit un possible contact entre la guanine 7 de sopC et le résidu R219 de SopB situé à l'extérieur du motif HTH de SopB.

Par ailleurs, des résidus impliqués dans l'interaction avec le centromère et situés en dehors du motif HTH, sont retrouvés chez plusieurs ParB plasmidiques. Par exemple, la fixation de ParB_{P1} sur son centromère est bipartite et se fait via son motif HTH et son domaine de dimérisation. Pour KorB du plasmide RK2, l'hélice de reconnaissance du motif HTH n'apporte pas la spécificité d'interaction. Ce sont les résidus Thr211 et Arg240, situés à l'extérieur de l'hélice de reconnaissance du motif HTH qui sont utilisés pour contacter spécifiquement le centromère (Khare *et al.*, 2004). Pour ParB de *P. aeruginosa*, les résidus G198 et R219 sont

également impliqués dans cette fonction (Kusiak et *al.* 2011). Dans le cas de SopB, le motif HTH a initialement été montré comme le seul déterminant impliqué dans l'interaction avec le centromère (Ravin et *al.*, 2003). Il nous a donc semblé intéressant de comprendre le rôle du résidu R219 dans le mécanisme de partition ainsi que dans l'interaction SopB-*sopC*.

Dans cette étude, nous avons montré que l'arginine (R219) située en dehors du motif HTH, est essentielle au processus de ségrégation car essentielle à l'interaction spécifique avec le centromère. Nous avons ensuite étendu notre étude à l'ensemble des ParB chromosomiques. En se basant sur des alignements de séquence et des expériences réalisées *in vitro* et *in vivo*, nous avons pu identifier et montrer qu'un résidu arginine des ParB chromosomiques, était aussi impliqué dans l'interaction spécifique avec le centromère. Ce résidu fait partie d'un motif très conservé et nouvellement identifié dans cette étude, le motif GHAR/K^A/VLL.

Ces résultats ont révélé une caractéristique conservée dans le règne bactérien : les protéines ParB contiennent un domaine de liaison au centromère, composé de deux motifs séparés et essentiels, le motif HTH et le motif GHAR/K^A/VLL.

II. Présentation de l'article: "Insight into centromere-binding properties of ParB"

Insight into centromere-binding properties of ParB proteins: a secondary binding motif is essential for bacterial genome maintenance

Aurore Sanchez^{1,2}, Jérôme Rech^{1,2}, Cyrielle Gasc^{1,2} and Jean-Yves Bouet^{1,2,*}

¹Laboratoire de Microbiologie et Génétique Moléculaires, Centre National de la Recherche Scientifique, F-31000 Toulouse, France and ²Laboratoire de Microbiologie et Génétique Moléculaires, Université de Toulouse, UPS, F-31000 Toulouse, France

Received October 1, 2012; Revised December 13, 2012; Accepted December 31, 2012

ABSTRACT

ParB proteins are one of the three essential components of partition systems that actively segregate bacterial chromosomes and plasmids. In binding to centromere sequences, ParB assembles as nucleoprotein structures called partition complexes. These assemblies are the substrates for the partitioning process that ensures DNA molecules are segregated to both sides of the cell. We recently identified the *sopC* centromere nucleotides required for binding to the ParB homologue of plasmid F, SopB. This analysis also suggested a role in *sopC* binding for an arginine residue, R219, located outside the helix-turn-helix (HTH) DNA-binding motif previously shown to be the only determinant for *sopC*-specific binding. Here, we demonstrated that the R219 residue is critical for SopB binding to *sopC* during partition. Mutating R219 to alanine or lysine abolished partition by preventing partition complex assembly. Thus, specificity of SopB binding relies on two distinct motifs, an HTH and an arginine residue, which define a split DNA-binding domain larger than previously thought. Bioinformatic analysis over a broad range of chromosomal ParBs generalized our findings with the identification of a non-HTH positively charged residue essential for partition and centromere binding, present in a newly identified highly conserved motif. We propose that ParB proteins possess two DNA-binding motifs that form an extended centromere-binding domain, providing high specificity.

INTRODUCTION

Low-copy number plasmids and most bacterial chromosomes have acquired partitioning (Par) mechanisms to

ensure faithful segregation in dividing cells. This stable inheritance depends on active partition systems [reviewed in (1)]. Virtually, all Par systems are composed of three essential elements: a centromere, a centromere-binding protein (CBP) and an NTPase. Interactions between these partners are important for the partition process, and of these, the specific recognition of centromeres by CBPs to assemble partition complexes is crucial. Partition complexes are substrates for partition machineries, as they are responsible for the proper intracellular localization of the replicon onto which these are assembled. CBPs are functional equivalents of eukaryotic kinetochores, in which they act as the link between centromeres and NTPases.

Bacterial partition systems have been classified in three types according to the nature of the NTPase involved (2). The partition system we study here is of type I, representative of the most widespread system, present on most plasmids and on all chromosomes carrying a Par system. Although, the molecular mechanisms underlying partitioning of type I systems are still to be unraveled, evidence is now accumulating that *in vivo* type I adenosine triphosphatases (ATPases) generally show dynamic behaviors, either oscillating over the nucleoid (3–5) or forming metastable filaments (6,7). Most of the ParA proteins have also been found to polymerize *in vitro* in response to adenosine triphosphate (ATP) binding (7–10). These dynamic behaviors depend on the presence of the three Par elements. Notably, ParB stimulates ParA activities, such as polymerization (8,9) and ATP hydrolysis (11,12). This latter activity is further stimulated with centromere-bound ParB (13). Partition complexes assemblies on centromere DNA are thus important for the regulation of the partition process.

Bacterial centromeres, generically called *parS* sites, are diverse in composition and organization. Chromosomal *parS* sites typically consist of 16-bp inverted repeats clustered around the replication origin or spread throughout the Ori domain (14). Plasmid centromeres consist of

*To whom correspondence should be addressed. Tel: +33 561 335 906; Fax: +33 561 335 886; Email: jean-yves.bouet@ibag.biotech.fr

© The Author(s) 2013. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

inverted and/or direct repeats [reviewed in (15)]. The repeats can be extensive, for example, *sopC* of the F plasmid consists of 12 tandem repeats of a 43-bp motif (16), and *parS* of pTAR contains 13 heptameric repeats that are separated by integral helical turns (17). By contrast, the partition site *O_B3* of plasmid RK2/RP4 is a simpler palindromic core sequence of 13 bp, although related sites dispersed throughout this plasmid might also contribute to segregation (18). Centromeres of more complex organization also exist, such as that of plasmid P1 (19). This ~80-bp *parS* consists of two arms that harbor non-symmetrical heptamer (A-box) and hexamer (B-box) DNA motifs, separated by a central binding site for the host-encoded DNA-binding protein, integration host factor. Despite this high diversity, ParBs generally bind to centromere sites through helix-turn-helix (HTH) or in some cases ribbon-helix-helix (RHH) motifs [for review (15)].

In F plasmid, the partitioning locus *sop* specifies the ATPase, SopA, the CBP, SopB and the centromere, *sopC*. Each of the 12 tandem 43-bp repeat contains a 16-bp inverted repeat that constitutes the SopB-binding site (20). SopB is composed of several functional domains (Figure 1A). Owing to its high flexibility, the structure of the N-terminal part of SopB is unknown, which is also the case with others ParBs (21,22). This part contains the SopA-interaction domain (23), the arginine-finger and the putative arginine-loop motifs that stimulate SopA ATPase activity (13) and possibly a SopB dimer-dimer interaction region (24). SopB not only forms a SopB-*sopC* partition complex that recruits SopA but also extends the complex by spreading along the DNA on either side of *sopC* (25), and also probably *in trans* to different DNA molecules (26). By coating DNA, SopB forms extended partition complexes whose role in partitioning is not understood. In all cases, specific assembly of SopB on *sopC* to nucleate the formation of extended partition complexes is essential for F partitioning.

Our present knowledge of the protein-centromere contacts that define binding specificity is based on structural studies of only three ParB-*parS* complexes: KorB-*O_B3* (RP4), ParB-*parS* (P1) and SopB-*sopC* (F). KorB binds *O_B3* through a classical HTH motif, and the recognition specificity is provided by two amino acid side chains external to the HTH motif (22). The P1 partition complex is more complicated owing to the unusual organization of its centromere site (see earlier in the text). Each HTH motif of ParB dimer contacts an A-box on each *parS* arm, and an additional DNA-binding module contacting the B-box is provided by the dimerization domain (21). The strong specificity of P1 ParB-*parS* interaction is shared between these two distinct DNA-binding motifs. Specificity determinants of SopB-*sopC* partition complex have been recently investigated at the nucleotide level by high-throughput analysis (27). This revealed that all 16 bp of *sopC* are required, to varying degrees, for maximal SopB binding strength. In addition, the co-crystal structure of a SopB-*sopC* complex was solved at ~3.5 Å (26), enabling prediction of the likely contacts between *sopC* and residues of the HTH motif (Figure 1C). Because previous results had suggested binding specificity was

entirely contained within the HTH (23), it was interesting that an arginine residue (SopB-R219) outside the HTH was close to the *sopC* guanine at position 7 (*sopC*-G7) in the co-crystal.

To better understand the underlying mechanisms involved in plasmid partitioning and centromere binding, we have now investigated the predicted interaction between SopB-R219 and *sopC*-G7. Strikingly, we found that this interaction is essential for SopB-*sopC* binding *in vitro* and is necessary for partition *in vivo*. These findings reveal that the specific contacts between SopB and *sopC* occur both inside and outside the HTH motif of SopB. Our results refine the basis of the high-binding specificity of F partition complex assembly. Based on sequence alignments and the known structure of a chromosomal ParB, we identified and showed that an arginine residue provides a specific contact with the chromosomal centromere. Extensive sequence analysis revealed that this positively charged residue is part of a newly defined and highly conserved motif present in chromosomal ParBs over the whole bacterial domain of life.

MATERIALS AND METHODS

Bacterial strains and plasmids

Strains are derivatives of *Escherichia coli* K-12 or *E. coli* B (C2833) and are listed, together with plasmids, in Supplementary Tables S1 and S2, respectively. Construction of plasmids is detailed in Supplementary Figure S4.

Media and growth conditions

Cultures were essentially grown at 37°C with aeration in LB (Miller, 1972) containing thymine (10 µg ml⁻¹) and antibiotics (Sigma) as appropriate (µg ml⁻¹): kanamycin (Km, 50), chloramphenicol (Cm, 10), streptomycin (Sm, 200) and spectinomycin (Sp, 20 in liquid medium, 30 in solid). Luria Broth (LB) was supplemented with 1.5% agar (Difco) for solid medium. For stability assay, cultures were grown at the indicated temperature with aeration in LB or in M9 minimal medium supplemented with glucose (0.4%), casamino acids (0.2%), MgSO₄ (1 mM), CaCl₂ (0.1 mM), thiamine (1 µg ml⁻¹), leucine (20 µg ml⁻¹) and thymine (40 µg ml⁻¹).

DNA manipulation and related procedures

Enzyme reactions, DNA preparation, agarose gel electrophoresis and transformation with plasmid DNA were carried out using standard procedures, according to suppliers' recommendations where applicable (restriction endonuclease, DNA polymerase and ligase, New England BioLabs; polymerase chain reaction with Phusion, Fermentas; plasmid and gel-fractionated fragment DNAs, Qiagen or Promega). SopB mutations were introduced in pYAS6 and pYAS47 plasmids by quick-change mutagenesis and verified by DNA sequencing (Supplementary Figure S4).

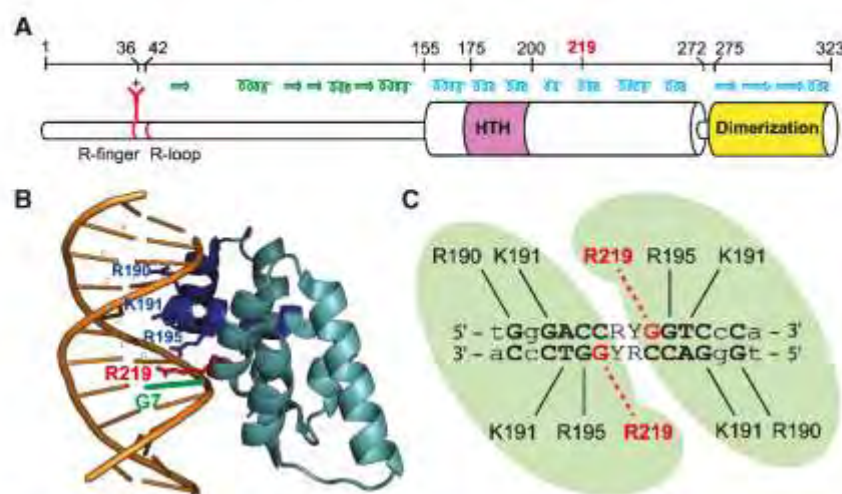
3096 *Nucleic Acids Research*, 2013, Vol. 41, No. 5


Figure 1. Model of SopB-*sopC* interaction. (A) Schematic representation of SopB. Top; amino acid numbering of SopB to scale. Middle; secondary structure of SopB. Helices and arrows represent α -helix and β -sheet, respectively, as determined by crystal analysis in blue or predicted (predator, JPRED) in green. Bottom; functional domains and motifs of SopB. Thick cylinders (from amino acids 155–272 and 275–323) represent the parts of SopB structure that have been solved, whereas thin cylinders (from amino acids 1–155 and 272–275) have not (26). (B) Representation of SopB-*sopC* structure drawn from accession number 3EZF (26), with R219 colored in red, R190, K191 and R195 in blue and the guanine 7 in green. Helices 2 and 3 of the HTH motif are colored in dark blue. (C) Predicted interactions between amino acids of SopB with *sopC* nucleotides; adapted from the study by Pillet *et al.* (27). R190, R219, R195 and K191 correspond to arginines 190, 195, 219 and lysine 191. Straight and dotted lines represent suggested amino acid-nucleotide interactions. R and Y stand for purine and pyrimidine nucleotides, respectively. Green zones represent the subunits of an SopB dimer.

Protein purification

Intein-tagged SopB-G324, SopB-R219 mutants, ParB_{Pput}-G291 and ParB_{Pput}-G291-R194A proteins were purified from strain C2833 essentially as described previously (13), except that fractions were directly collected and frozen at -80°C after elution from the chitin column. ParB_{Bcen}-G306 does not bind to the chitin column. ParB_{Bcen}-G306 and ParB_{Bcen}-G306-R192A were prepared as described earlier in the text, except that Frl was incubated directly with dithiothreitol to induce full intein cleavage (see Supplementary Figure S3) and was stored in aliquots at -80°C . ParB proteins from *Bcen* were assayed directly with crude extracts.

Plasmids stability assay

Experiments were started from colonies of cells freshly transformed with the plasmids under test. SopB alleles, expressed from a pSC101 derivative under *plac*, were assayed in strain MC1061. Overnight cultures in M9-glucose-casamino acid medium containing selective antibiotics were diluted 400-fold into the same medium and grown to $A_{600} = 0.25$. Samples were then diluted serially into fresh M9 medium without the antibiotic selective for the plasmid under test and were processed as described previously (28). To determine the fraction of cells that retained the plasmid, samples were taken at the beginning and after 20 and 40 generations. The calculated loss rate per generation is given by the slope A of the equation ($y = B \cdot e^{-Ax}$) from the fitted curves (using Graph Pad Prism 4 software) on graph plotting the per cent of cell mini-F⁺ as a function of the generation

time, which provides a good approximation of the loss frequency (f) when $<10\%$ [$f = 1 - (\text{cell mini-F}^+ / \text{total cell})^{1/n}$, where n is the generation number]. ParB alleles of *Bcen* or *Pput* were assayed for the stabilization of pDAG504 in same conditions as previously published (29,30). Briefly, for *Bcen* alleles, plasmids under test were introduced in strain DH10B, and growth cultures were performed in M9-glucose-casamino acid medium at 37°C . For *Pput* alleles, plasmids under test were introduced into strain DH5 α , and growth cultures were performed in LB at 37°C .

Electromobility shift assay

Experiments were performed as previously described (9), using 1 nM radiolabeled DNA probes in the presence of sonicated salmon sperm DNA as competitor ($100 \mu\text{g ml}^{-1}$). Binding reaction products were resolved at 4°C by electrophoresis on 6% polyacrylamide gels in TGE (25 mM Tris base, 190 mM Glycine, 1 mM Ethylenediaminetetraacetic acid) for 2–4 h at 180 V. Polyacrylamide gel electrophoresis-purified complementary oligonucleotides (Sigma-Aldrich) were used to assemble duplex DNA probes by annealing (only top strands are shown; 16-bp SopB/ParB binding sites are indicated by upper case letters; bases corresponding to positions 7 and 10 are indicated in bold): sopC2: 5'-attagtcTGGGACCACGGT CCCActegtat-3'; sopC63: 5'-attagtcTGGGACAACCTGTC CCActegtat-3'; sopC64: 5'-attagtcTGGGACTACAGTCC CActegtat-3'; parS66: 5'-attagtcTGTTCCACGTGGAAC Actegtat-3'; and parS67: 5'-attagtcTGTTTCACGTGAAA CActegtat.

RESULTS

We have previously investigated the bases involved in the specific recognition of the F plasmid centromere by SopB (27); summarized in Figure 1C]. In particular, the guanine at position 7 (*sopC*-G7) in the left arm (and its counterpart at position 10 in the right arm) was found to play a role in SopB *sopC* interaction. In addition, the co-crystal structure of SopB-*sopC* complexes, determined at a medium resolution [~ 3 Å; (26)], suggests that G7 is close to an arginine residue (R219), in the major groove (Figure 1B). We estimated the distance between the O group of G7 and the NH group of R219 at 2.6 Å (Supplementary Figure S1), compatible with hydrogen-bonding between them (31). This R219 residue does not belong to the HTH motif responsible for the primary binding to DNA (Figure 1A), which contains three positively charged residues (R190, K191 and R195; Figure 1B) involved in specificity of binding to *sopC* (13,26). SopB-R219 is part of an α -helix separated from the HTH by another α -helix (Figure 1A). Its potential role in centromere binding and thus in the partition process was further investigated.

SopB-R219 mutants are unable to stabilize mini-F

To determine whether the R219 residue plays a direct role in the partition of the mini-F plasmid, we constructed two SopB-R219 mutants by converting R219 to alanine (R219A) or lysine (R219K). These mutations were introduced into a pSC101 derivative, pYAS47, which carries *sopAB* under the control of the *lac* promoter, yielding pJYB161 (*sopB*-R219A) and pJYB162 (*sopB*-R219K). These plasmids were introduced into strain MC1061 carrying pDAG209, an unstable mini-F *sopC*⁺ but Δ *sopAB*, and loss of the mini-F during growth without selection was measured (Supplementary Figure S2A and Table 1). When wild-type SopB was provided *in trans* from pYAS47, no cell without pDAG209 was observed after 40 generations, indicating that the mini-F *sopC*⁺ was completely stable (loss rate $< 0.01\%$). Without SopB (pAM238), the mini-F was lost at a frequency of 4.0% per generation as expected from the copy number ($\sim 5-6$ at division) for segregation at random. The mutant alleles SopB-R219A (pJYB161) and -R219K (pJYB162) failed to stabilize pDAG209, which was lost at 3.4 and 3.1% per generation, respectively. Western blot analysis showed that in the stability assays, the mutant SopBs were produced at the same level as wild-type (wt) (Supplementary Figure S2B). These results indicate that the R219 residue of SopB is essential for the maintenance of mini-F during generation.

SopB-R219 variants do not bind *sopC* centromere

To test whether the inability of *sopB*-R219 alleles in stabilizing the mini-F is primarily because of a default in centromere-binding activity, we performed electromobility shift assays (EMSA). Wild-type SopB and R219 variants were expressed as intein-tagged proteins and purified as native proteins after cleavage of the intein (see 'Materials and Methods' section and Supplementary Figure S3).

Typical EMSA experiments are shown in Figure 2A. The ³²P-labeled substrate, a 30-bp DNA duplex containing the 16-bp consensus SopB-binding site, was specifically bound by SopB as previously observed (27). By contrast, SopB-R219A showed almost no *sopC*-binding activity: at 1 μ M, the amount of protein-DNA complex formed is much less than with 10 nM of wild-type SopB, indicating that the loss of *sopC*-binding activity in SopB-R219A is >200 -fold. The arginine 219 of SopB is, therefore, essential for specific SopB binding to the centromeric site. Similarly, SopB-R219K showed a significant reduction in *sopC*-binding activity, although slightly less than SopB-R219A. The ~ 100 -fold loss in *sopC* binding indicates that the lysine present in this variant is not sufficient to compensate for the positive charge of the arginine residue in wild-type SopB.

In addition to its site-specific DNA-binding activity, SopB also binds non-specific DNA (9), for spreading along the DNA outside of *sopC* (25). To determine the extent to which SopB-R219A variant is deficient in DNA-binding activity, we also tested its ability to still bind DNA non-specifically. For this, we performed EMSA in the absence of competitor DNA, using a 114-bp duplex DNA that does not carry the 16-bp *sopC* site (Figure 2B). In these conditions, SopB-wt and SopB-R219A bound similarly the non-specific DNA probe. For both alleles, one discrete complex (NS1) was observed between 100 and 200 nM, and a second discrete complex (NS2) was detected between 200 and 500 nM. At ~ 500 nM SopB, $\sim 50\%$ of the probe was shifted, and above this concentration, high molecular weight complexes were formed. Thus, both proteins displayed about the same affinity for non-specific DNA, indicating that SopB-R219 is deficient in *sopC* binding only. This result strengthens the finding that the R219 residue of SopB is strongly involved in site-specific DNA binding.

In vivo SopB-R219A does not form foci

SopB, like all ParB proteins, forms bright foci in cells carrying its cognate centromere (32). These foci correspond to the assembly of partition complexes on centromeres and are usually located near the cell center or the quarter positions (33). To determine whether SopB binding to *sopC* was also affected *in vivo* by the alanine mutation at position 219, we constructed mini-F derivatives, pJYB214 and pJYB223, with Venus-Yfp fused to SopB-wt and SopB-R219A, respectively. pJYB214 was fully stable in strain DLT1215 (data not shown), indicating that the fusion is functional. In this strain, DLT2950, fluorescence microscopy showed that all cells harbor two to four bright regularly positioned foci (Figure 3, left panel). By contrast, cells transformed with pJYB223, which carries the *sopB*-R219A allele, did not form foci, and fluorescence was spread throughout the cytoplasm (Figure 3, right panels). In addition, some cells ($\sim 15\%$) did not display fluorescence, suggesting that they had lost the mini-F, as expected for an unstable mini-F derivative. The absence of focus formation implies that SopB-R219 does not support centromere binding *in vivo*.

Table 1. Stability of mini-F plasmids carrying F *sopC* or chromosomal *parS* centromeres

| F plasmids | Multicopy plasmids | Relevant characteristics | Loss/generation (%) |
|--|--|--|---|
| pDAG209 (<i>sopC</i>) | pAM238 pYAS47 pJYB161 pJYB162 | <i>sopA</i> ⁺ <i>B</i> ⁺ <i>sopA</i> ⁺ <i>B-R219A</i> <i>sopA</i> ⁺ <i>B-R219K</i> | 4.0 (±0.5) <0.01 3.4 (±0.3) 3.1 (±0.1) |
| pDAG551 (<i>parS</i>) | pBBR1MCS5 pDAG562 pJYB167 | <i>Bcen parA</i> ⁺ <i>B</i> ⁺ <i>Bcen parA</i> ⁺ <i>B-R192A</i> | 2.6 (±0.1) 0.15 (±0.1) 2.7 (±0.4) |
| pDAG504 (<i>parS_{ppu}</i>) | pAM238 pDAG503 pJYB168 | <i>Pput parA</i> ⁺ <i>B</i> ⁺ <i>Pput parA</i> ⁺ <i>B-R194A</i> | 3.6 (±0.7) 0.25 (±0.1) 2.1 (±0.5) |

Derivatives of strain MC1061 carrying the mini-F pDAG209 and the indicated plasmids were grown at 30°C. Strains carrying the mini-F's pDAG551 (consensus *parS*) or pDAG504 (*parS_{ppu}*) were grown at 37°C. The *sop* and *par* operons are under the control of the *lac* promoter on pAM238 or pBBR1MCS5. Samples of the cultures were plated at intervals for estimation of the proportion of mini-F-positive cells. Loss rates per generation, given by the slopes (A) of the fitted curves as presented in Supplementary Figure S2A, are average of at least three experiments, and standard deviations are indicated in brackets.

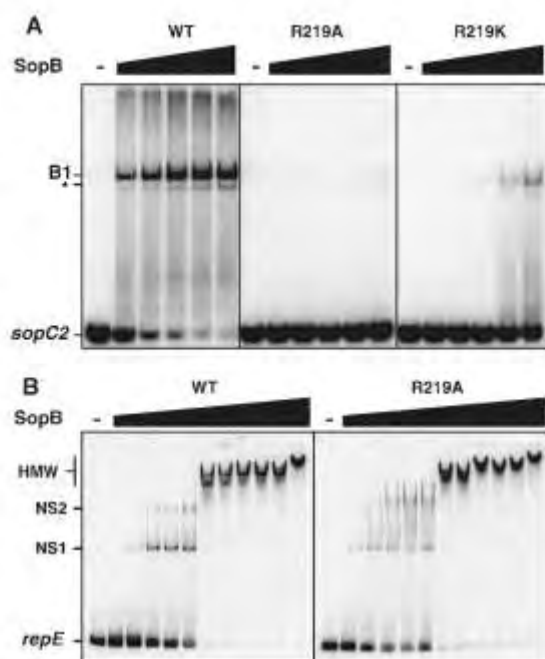


Figure 2. Interaction of SopB wild-type and R219 variants with DNA. The ability of wild-type SopB, SopB-R219A and -R219K to interact with an *sopC* DNA duplex or with non-specific DNA was tested in a gel mobility shift assay. Reaction mixtures, assembled on ice and incubated at 30°C, were analyzed by electrophoresis on polyacrylamide gels. (A) SopBs binding to *sopC* centromere site. ³²P-labeled 30-bp *sopC2* probe was incubated alone (–) or with increasing concentrations (from left to right; 10, 30, 100, 300 and 1000 nM) of wild-type SopB (WT), SopB-R219A or SopB-R219K as indicated on top, in the presence of 100 µg ml^{–1} competitor DNA. Positions of DNA probes and protein complexes are indicated on the left: B1 denotes the specific complex, whereas asterisk corresponds to a specific shift from a secondary form of the duplex DNA. (B) SopBs binding to non-specific DNA probe. ³²P-labeled 114-bp *repE* probe was incubated alone (–) or with increasing concentrations (from left to right; 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 and 2000 nM) of wild-type SopB (WT) or SopB-R219A as indicated on top, in the absence of competitor DNA. NS1 and NS2 denote discrete complexes, and HMW corresponds to high molecular weight complexes.

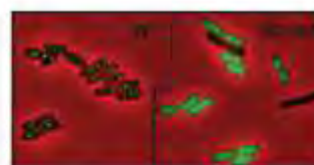


Figure 3. SopB-R219A allele does not form foci cells. Left: localization of SopB-Venus fusion. Fluorescent foci from the mini-F expressing SopB-Venus display the expected pattern for partition complexes faithfully partitioned. Right: localization of SopB-R219A-Venus fusion. All cells displayed diffuse fluorescence throughout the cytoplasm. Cells without fluorescence are presumed to lack the mF carrying SopB-R219A-Venus as expected from the loss rate (see Supplementary Figure S2A and main text).

Guanines 7 and 10 in *sopC* site are critical for SopB binding

The SopB-*sopC* co-crystal structure (26) predicts that the R219 residue of an SopB monomer is near the guanine in position 7 of the left arm of *sopC*, and thus that R219 of the second SopB monomer should be close to the guanine 10 on the right arm (Figure 1B and C). We previously observed by SPRi that mutations of these bases affect SopB binding to *sopC* (27). However, because of the low sensitivity of this technique, quantification of the loss in DNA-binding activity was not possible. Because SopB-R219 variants are deficient in *sopC* binding, we expected that if R219 directly interacts with guanines 7 and 10, we should observe a comparable loss in SopB binding if these bases are mutated.

We, therefore, performed EMSA in same conditions as described earlier in the text with wild-type SopB but with 30-bp *sopC* DNA probes carrying either transversion (C to A; probe *sopC63*) or transition (C to T; probe *sopC64*) at positions 7 and 10 (Figure 4). Both probes were barely shifted even at the highest SopB concentration (1 µM), indicating that these bases are strictly required for the formation of SopB-*sopC* complexes. This result is consistent with a direct interaction between R219 residues and guanine at positions 7 and 10.

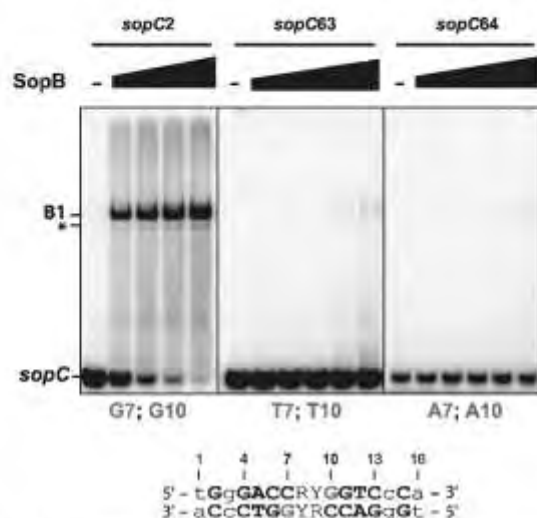


Figure 4. Interaction of wild-type SopB with mutated *sopC* centromeric DNA. Top: the ability of wild-type SopB to interact with different *sopC* DNA duplexes (indicated on top) was assayed as in Figure 2. 32 P-labeled 30-bp *sopC* probes were incubated alone (–) or with increasing concentrations of wild-type SopB (from left to right; 10, 30, 100, 300 and 1000 nM), except for *sopC2* where the 1000 nM SopB reaction was omitted. Legend is as in Figure 2. Bottom: bases at positions 7 (bottom strand) and 10 (top strand) are indicated in grey. Numbers indicate positions of the bases in the palindromic SopB-binding site according to the study by Pillet *et al.* (27).

Chromosomal CBPs harbors extended DNA-binding domains essential for partition activity

The structure of a chromosomal homologue of SopB, Spo0J of *Thermus thermophilus*, solved in the absence of the *parS* centromeric DNA, also displays a typical HTH proposed to be the *parS*-binding motif (34). Interestingly, a helix (H9) located two helices away from the HTH recognition helix (H7) was shown to correspond to helix H8 of KoriB (35), and it also aligns with helix H5 of SopB (Figure 6A). H5_{SopB} and H8_{KoriB} carry the R219 and the R240 residues, respectively, involved in specific recognition of their cognate centromere [this study, (22)]. Helix H9_{Spo0J} also harbors an arginine residue (R179) in the corresponding helix (Figure 6A). In the 3D structure of the Spo0J dimer, the orientation of R179 relative to the HTH supports the possibility that it could provide a direct interaction with a specific base of *parS* (Figure 6B). Indeed, each arginine residue on both monomers has its side chain directed towards where the *parS* DNA would be bound in the HTH motif alignment.

Sop0J_{the} binds the pseudopalindromic *parS* sequence 5'-TGTTCCACGTGAAACA-3'. This sequence corresponds to the highly conserved *parS*, characteristic of centromeres present on bacterial chromosomes, whose consensus is 5'-tGTTCACGTGNAACA-3'. If the Spo0J-R179 residue is involved in specific binding, we expect that it should be conserved at least in a subset of chromosomal ParBs. To test this hypothesis, we aligned ParBs from a large range of bacteria (Figure 6C). We found that R179 is highly conserved and present in most

phyla. In the ϵ -Proteobacteria, the Chlamydiales and the Spirochaete groups, the arginine is replaced by a lysine, suggesting that the positively charged nature of the residue is the important feature, compatible with a direct interaction with a base of the *parS* site. If the H9 helices of ParB are functional equivalents to the H5 helix of SopB, these R/K residues should also be involved in chromosome segregation.

ParABS partition activity on bacterial chromosome is difficult to be assessed directly because of pleiotropic effects. However, partition function could be examined by the ability of *parAB* expressed from *plac* on a multicopy vector to stabilize a mini-F (Δ *sop*) plasmid carrying a single *parS* in *E. coli* *in trans*. Indeed, this assay has been used previously to demonstrate the partition capacity of the *Bacillus subtilis* *soj-spo0J-parS* and the *Pseudomonas putida* and *Burkholderia cenocepacia* *parABS* systems (29,30,36). To determine whether these positively charged residues are involved in partition, we tested the ParABS_{*Pput*} and ParABS_{*Bcen*} systems in stability assays by introducing mutations at positions equivalent to Spo0J-R179, *parB*-R194A and *parB*-R192A, respectively. We observed, as previously reported using the same experimental conditions (29,30), that both ParAB systems were able to efficiently stabilize their cognate *parS* carrying mF, pDAG504 or pDAG551 (Table 1; loss rates = 0.15–0.25%). We found, however, that both *parB*-R194A and *parB*-R192A alleles were impaired in their ability to support stabilization of pDAG504; *Bcen* ParB-R192A was completely deficient (loss rate = 2.7%), whereas *Pput* ParB-R194A was highly but not completely deficient (loss rate = 2.1%). These results thus clearly demonstrate that these positively charged residues in H9 helices are essential, as is SopB-R219, for the partition function of chromosomal ParBs.

By performing EMSA, we then tested whether these arginine residues are, as for their SopB counterparts, directly involved in *parS*-specific binding (Figure 5). ParB proteins from *P. putida* were purified as for SopBs, whereas ParB proteins from *B. cenocepacia* were directly assayed from crude extracts expressing intein-tagged ParBs (see 'Material and Methods' section and Supplementary Figure S3). In contrast to wild-type, mutant ParBs from *Pput* and *Bcen* were not able to bind to their respective *parS* sites. In both cases, the DNA binding deficiency is so high that no specific retardation was detected. These results thus indicate that R194 and R192 residues from *Pput* and *Bcen*, respectively, are directly involved in centromere binding.

Finally, we found that the R/K residues of H9 helices are part of a highly conserved motif, GHA^{R/K}/V^ALL, that covered most of the H9 helix (Figure 6C). Secondary structure analysis on these motifs predicted that they all harbor an α -helical conformation (data not shown). We named this conserved motif box 'CBM2' for centromere-binding motif 2; the CBM1 being the typical HTH motif.

DISCUSSION

SopB is known to bind DNA specifically on the *sopC* centromere to function in the active partitioning of the

3100 Nucleic Acids Research, 2013, Vol. 41, No. 5

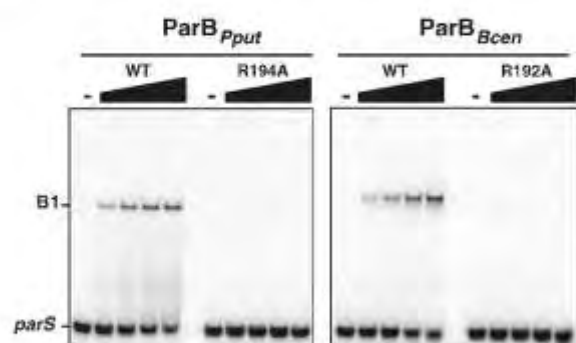


Figure 5. Chromosomal ParB binding to *parS* sites. DNA binding interaction between *Pput* or *Bcen* *parS*, left and right panel, respectively, and wild-type ParB from *Pput* (*ParB_{Pput}*) or *Bcen* (*ParB_{Bcen}*) or variant *ParB-R192A_{Bcen}* and *ParB-R194A_{Pput}* proteins were assayed by EMSA as in Figure 2. ³²P-labeled 30-bp *parS* probes were incubated alone (–) or with increasing concentrations of purified *ParB_{Pput}* proteins (left panel; from left to right: 0, 10, 30, 100 and 300 nM) or *ParB_{Bcen}* proteins in crude extracts (right panel; from left to right: 0, 2.8, 8.4, 28 and 84 mg ml^{–1}). Legend is as in Figure 2.

F plasmid. We have shown here that the specificity of SopB *sopC* interaction is more complex than anticipated, by identifying a new specific DNA-binding determinant, an arginine at position 219, located outside of the HTH motif. We demonstrated that SopB-R219A/K variants are completely deficient in *sopC* binding *in vitro*, whereas both mutants are still able to bind DNA non-specifically (Figure 2). Fluorescence microscopy revealed that, in contrast to wt SopB, neither mutant allele was able to form intracellular foci (Figure 4), supporting the idea that *in vivo* they are also deficient in *sopC* binding. In addition, these single mutations totally abolished active segregation (Table 1), demonstrating that the R219 residue is a crucial DNA-binding determinant of the SopB–*sopC* interaction. Extension of this result to chromosomal partition systems reveals that most ParBs also harbor a secondary CBM that was shown to be essential for partitioning functions (Figure 6 and Table 1).

Direct interaction between R219 and guanine 7 defines an extended DNA-binding domain in F SopB

HTH motifs are composed of two helices that contact the DNA in the major groove. The second helix, called the recognition helix, usually provides specificity by contacting specific bases of the target binding sequence. In the case of SopB, the recognition helix makes three specific base contacts through positively charged residues (K190, K191 and K195) and possibly two others through residues S189 and I192 (26). These numerous contacts were thought to be sufficient for providing a high specificity for centromere binding. The proximity evident in the SopB–*sopC* structure of the positively charged residue, R219, located two helices away from the HTH, to guanine at position 7 (G7) was thus intriguing, especially because G7 was found to be involved in SopB binding (27). We found the same strong deficiency of binding, as judged by gel retardation, between

SopB-R219A binding to *sopC* and wild-type SopB binding to *sopC* carrying either transitions or transversions at positions 7 and 10 (compare Figures 2A and 4). In addition, replacing the arginine by a lysine at position 219 does not restore binding to *sopC* (Figure 2A). A positive charge at this position is thus not sufficient for DNA binding, arguing for specific contact between SopB-R219 and *sopC*-G7. In agreement with this finding, bioinformatic analysis (27) showed that (i) the arginine at position 219 is conserved among all closely related SopB homologues; and (ii) all *sopC* centromere sequences present in the databases harbor a guanine at position 7. Along with these data, our results thus clearly support a direct interaction between R219 residues of each SopB monomer with each G7 base of the palindromic *sopC* sequence.

If the R219-G7 interaction was anticipated from the structure, the finding that it is a critical one was not. Single mutations at position R219 completely abolish *sopC* recognition, revealing that the HTH is not sufficient to provide the specificity of the SopB–*sopC* interaction. However, specific contacts made by the recognition helix are also crucial for binding specificity, as mutations of two of the positively charged residues, K191 and K195, also abolish specific, but not non-specific, *sopC* binding (13). The centromere-binding specificity of plasmid F is thus shared by two motifs, the HTH and the R219 containing helix, which together define an extended DNA-binding domain.

Centromere-binding proteins harbor extended DNA-binding domains for highly specific interactions

Is this extended centromere-binding domain present in SopB a common characteristic among CBPs? Some co-crystal structures of ParB–*parS* are available to address such a question. Similarly to SopB, KorB of RK2/RP4 harbors two DNA-binding motifs (22). However, in sharp contrast to SopB, residues from the helices forming the HTH motif of KorB do not make specific contacts with bases in the O_B site (13-bp *parS* centromere). The KorB–O_B recognition specificity is mainly based on two side chain interactions outside the HTH (T211 and R240) that form direct hydrogen bonds with two bases next to the central base pair binding sequence. Thus, although the specificity determinants are not distributed as for F SopB, KorB also has an extended centromere-binding domain.

In the case of the P1 ParB, the specificity recognition is shared by two separate DNA-binding motifs, each contacting its specific binding sequence, the A- and B-boxes; the P1 centromere site being composed of a complex arrangement of these two DNA-binding boxes (37,38). The A-boxes are specifically bound by the HTH motifs, and the B-boxes are bound by the dimerized dimer motifs (21). These two motifs participate to the efficient binding of ParB to *parS*; therefore, they could be regarded as an extended DNA-binding domain. Although P1 ParB and F SopB are closely related to each other, SopB does not carry this second DNA-binding motif because of the lack of two small extra loops residues providing ‘wings’

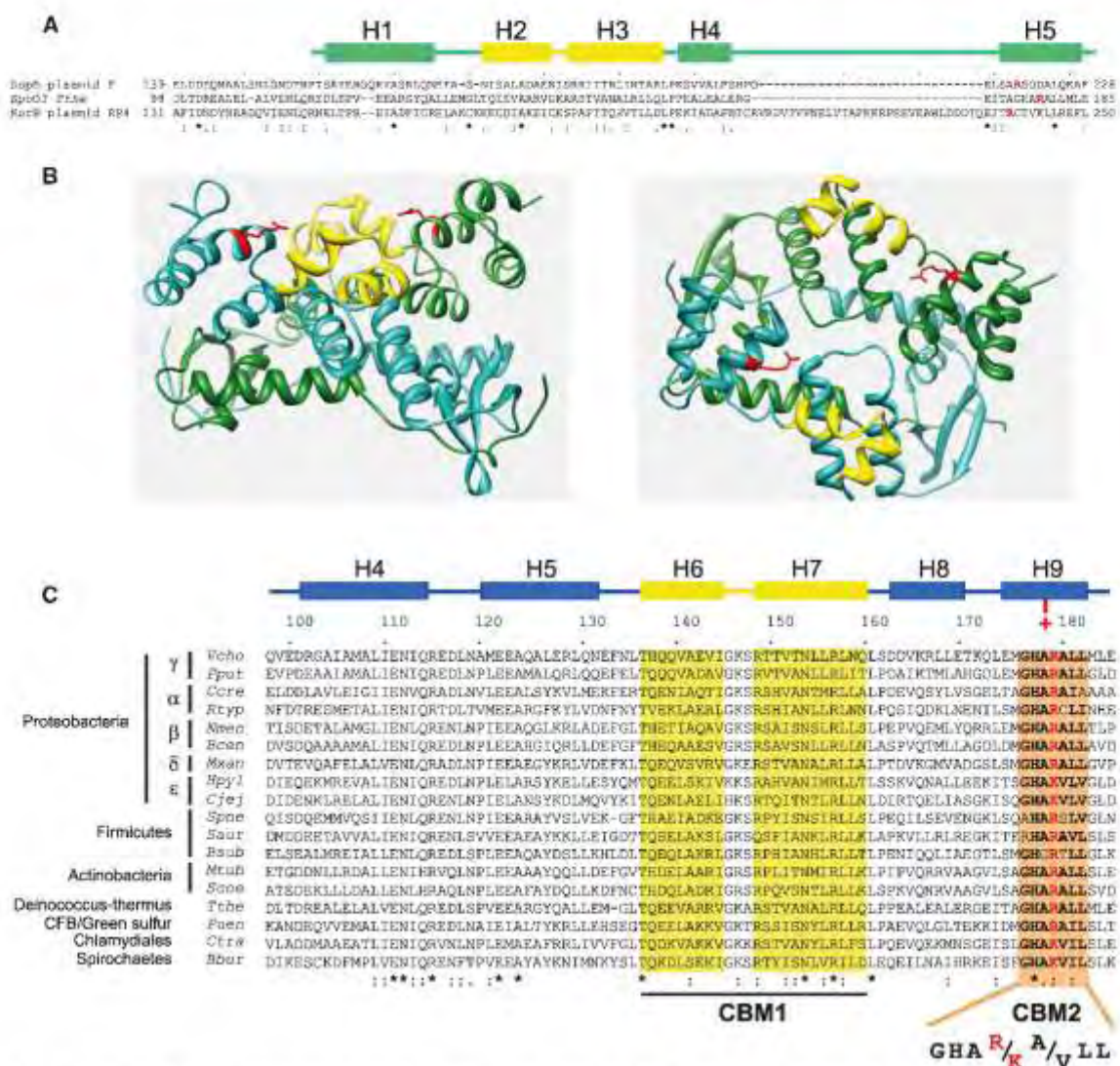


Figure 6. ParB conserved motifs and residues involved in centromere binding. (A) Amino acid alignment of ParB homologues. Sequence alignments of SopB (F) with Korf (RP4) and Spo0J (*T. thermophilus*) are based on the study by Leonard *et al.* (34). Identical residues and highly conserved residues are indicated by asterisks and dots, respectively. Helices H1 to H5 of SopB, based on the crystal structure analysis (26), are schematically represented on top; Helices H2 and H3, corresponding to the HTH motif, are drawn in yellow. Arginine residues R219 (SopB), R240 (Korf) and R179 (Spo0J) are colored in red. (B) Ribbon diagrams of a ParB dimer. Structure is from *T. thermophilus* Spo0J (34). One monomer is colored in cyan and the other in green. The right diagram is a 90° horizontal rotation of the left diagram. HTH motifs and R179 residues of each monomer are colored in yellow and red, respectively. Each arm of the 16-bp *parS* DNA is expected to bind one HTH motif of the Spo0J dimer. Orientations of R179 residues are compatible for providing additional specificity contacts in *parS* DNA. (C) Amino acid alignment of chromosomal ParBs. ParBs from a large spectrum of bacterial species were aligned with ClustalW; in the case of multichromosomal bacteria, ParBs are from the main chromosome. These ParBs are expected to bind the *parS* consensus sequence 5'-GTTNACGCTGNAAC-3' (30,48). Sequences highlighted in yellow and orange correspond to the HTH motif and the CBM2 (centromere-binding motif 2) box, respectively. Helices are represented as in (A) based on the crystal structure of Spo0J of *T. thermophilus* (34); amino acid sequence are numbered according to *T. thermophilus*. The red plus sign (+) schematically represents the conserved R or K positively charged residues. ParBs are from *Borrelia burgdorferi* B31 (Bbur), *B. cenocepacia* (Bcen), *B. subtilis* (Bsub), *Campylobacter jejuni* (Cfej), *Cautioner crescentus* (Ccre), *Chlamydia trachomatis* (Ctra), *Helicobacter pylori* (Hpyl), *Mycobacterium tuberculosis* (MtuB), *Mycoplasma xanthus* (Mxan), *Neisseria meningitidis* (Nme), *P. putida* (Pput), *Porphyromonas uenonis* (Puen), *Rickettsia typhi* (Rtyp), *Staphylococcus aureus* (Saur), *Streptomyces coelicolor* (Scoe), *Streptococcus pneumoniae* (Spne), *T. thermophilus* (Tthe) and *Vibrio cholera* (Vcho).

that are specifically present in the ParB dimerization motif (26). This difference is in relation with the centromere structures where, in the particular case of P1, ParB recognizes two different binding sequences. Thus, as in

Korf and SopB, P1 ParB has an extended DNA-binding domain composed of two centromere-binding motifs.

In the course of this study, we have identified the SopB-R219 counterpart present on chromosomal ParBs in most

bacterial phyla (Figure 6C). As for SopB, this residue present in the two chromosomal Par systems tested here is essential for the partition function (Table 1) and is absolutely required for binding to its cognate centromere sequence (Figure 5). In addition, it belongs to a newly identified and well-conserved α -helical motif, GHAR_K^A/VLL. This motif aligns with helices H5 and H8 of SopB and Korb, respectively, which are also essential for centromere binding. We, therefore, named this new motif CBM2 for centromere-binding motif 2. We proposed that chromosomal ParBs also harbor an extended centromere-binding domain composed of two DNA-binding motifs, the classical HTH (CBM1) and the newly identified CBM2.

We were unable to identify the CBM2 box and the positively charged residue on annotated chromosomal ParBs in only a few phyla, e.g. in cyanobacteria. The CBM2 motif, GHAR_K^A/VLL, is also not conserved on ParBs from secondary chromosomes in the case of bacteria carrying two or three chromosomes (data not shown). This could be because of divergent *parS* sites present on these chromosomes. Indeed, in the case of secondary chromosomes, *parS* sequences are different from the consensus one present on the main chromosome (14,39). It will be of interest to investigate whether a second DNA-binding motif is present on these Par systems.

In the case of type Ib ParBs that carry RHH DNA-binding motifs in place of an HTH, no co-crystal structure is available to determine whether specificity determinants are also present outside the primary binding motif. However, further DNA contacts providing additional binding specificity have been suggested by comparison of ParG, the ParB analogue of TP228, with the Arc and Mnt repressors harboring RHH motifs (40), supporting the possibility that type Ib ParBs could also carry extended centromere-binding domains.

Initially, the HTH motif of SopB was identified as the only determinant for centromere-specific binding (23). Our results along with these examples indicate that ParB-*parS* interactions are not as simple as initially proposed. Rather they suggest that bacterial CBPs have evolved, using different strategies, to possess extended DNA-binding domains. We thus propose that the requirement for at least two independent DNA-binding motifs is imposed by the need for a high centromere-binding specificity. Mutational analysis and co-crystal structure determination of other ParB-*parS* complexes, especially from chromosomal Par systems but also from ParBs with RHH, will be of particular interest to determine the extent of the different strategies developed for providing such high specificity for centromere binding.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1 and 2, Supplementary Figures 1–4 and Supplementary References [13,29,30,41–47].

ACKNOWLEDGEMENTS

The authors thank S. Depaul for technical assistance with stability assays and media preparation, N. Seddiki for help with PyMol software, Y. Ah-Seng for providing pYAS47 before publication, V. Morales for her expertise in protein purification and all members of the ‘segregation’ team for discussions. They are grateful to C. Johnson and reviewers for improvements in the manuscript.

FUNDING

Agence National pour la Recherche [2010 BLAN 1316 01]; Fonds Europ  en de D  veloppement Regional (for microscope equipments); Association pour le Recherche sur le Cancer (for microscope equipments). Funding for open access charge: Agence National pour la Recherche [2010 BLAN 1316 01].

Conflict of interest statement. None declared.

REFERENCES

- Salje, J. (2010) Plasmid segregation: how to survive as an extra piece of DNA. *Crit. Rev. Biochem. Mol. Biol.*, **45**, 296–317.
- Gerdes, K., M  ller-Jensen, J. and Bugge-Jensen, R. (2000) Plasmid and chromosome partitioning: surprises from phylogeny. *Mol. Microbiol.*, **37**, 455–466.
- Fogel, M.A. and Waldor, M.K. (2006) A dynamic, mitotic-like mechanism for bacterial chromosome segregation. *Genes Dev.*, **20**, 3269–3282.
- Ebersbach, G. and Gerdes, K. (2001) The double par locus of virulence factor pB171: DNA segregation is correlated with oscillation of ParA. *Proc. Natl Acad. Sci. U.S.A.*, **98**, 15078–15083.
- Marston, A.L. and Errington, J. (1999) Dynamic movement of the ParA-like Soj protein of *B. subtilis* and its dual role in nucleoid organization and developmental regulation. *Mol. Cell*, **4**, 673–682.
- Jakimowicz, D., Zydek, P., Kois, A., Zakrzewska-Czerwinska, J. and Chater, K.F. (2007) Alignment of multiple chromosomes along helical ParA scaffolding in sporulating *Streptomyces hyphae*. *Mol. Microbiol.*, **65**, 625–641.
- Ptacin, J.L., Lee, S.F., Garner, E.C., Toro, E., Eckart, M., Comolli, L.R., Moerner, W.E. and Shapiro, L. (2010) A spindle-like apparatus guides bacterial chromosome segregation. *Nat. Cell Biol.*, **12**, 791–798.
- Barilla, D., Rosenberg, M.F., Nobbmann, U. and Hayes, F. (2005) Bacterial DNA segregation dynamics mediated by the polymerizing protein ParF. *EMBO J.*, **24**, 1453–1464.
- Beut, J.Y., Ah-Seng, Y., Benmeradi, N. and Lane, D. (2007) Polymerization of SopA partition ATPase: regulation by DNA binding and SopB. *Mol. Microbiol.*, **63**, 468–481.
- Hui, M.P., Galkin, V.E., Yu, X., Stasiak, A.Z., Stasiak, A., Waldor, M.K. and Egelman, E.H. (2010) ParA2, a *Vibrio cholerae* chromosome partitioning protein, forms left-handed helical filaments on DNA. *Proc. Natl Acad. Sci. U.S.A.*, **107**, 4590–4595.
- Davis, M.A., Martin, K.A. and Austin, S.J. (1992) Biochemical activities of the ParA partition protein of the P1 plasmid. *Mol. Microbiol.*, **6**, 1141–1147.
- Watanabe, E., Wachi, M., Yamasaki, M. and Nagai, K. (1992) ATPase activity of SopA, a protein essential for active partitioning of F-plasmid. *Mol. Gen. Genet.*, **234**, 346–352.
- Ah-Seng, Y., Lopez, F., Pasta, F., Lane, D. and Bouet, J.Y. (2009) Dual role of DNA in regulating ATP hydrolysis by the SopA partition protein. *J. Biol. Chem.*, **284**, 30067–30075.
- Livny, J., Yamaichi, Y. and Waldor, M.K. (2007) Distribution of centromere-like *parS* sites in bacteria: insights from comparative genomics. *J. Bacteriol.*, **189**, 8693–8703.

15. Hayes, F. and Barilla, D. (2006) The bacterial segrosome: a dynamic nucleoprotein machine for DNA trafficking and segregation. *Nat. Rev. Microbiol.*, **4**, 133–143.
16. Hilsberg, M. and Eichenlaub, R. (1986) Twelve 43-base-pair repeats map in a *cis*-acting region essential for partition of plasmid mini-F. *J. Bacteriol.*, **165**, 1043–1045.
17. Kalnin, K., Stegalkina, S. and Yarmolinsky, M. (2000) pTAR-encoded proteins in plasmid partitioning. *J. Bacteriol.*, **182**, 1889–1894.
18. Williams, D.R., Macartney, D.P. and Thomas, C.M. (1998) The partitioning activity of the RK2 central control region requires only *incC*, *korB* and *KorB*-binding site *O*(B)3 but other *KorB*-binding sites form destabilizing complexes in the absence of *O*(B)3. *Microbiology*, **144**, 3369–3378.
19. Martin, K.A., Davis, M.A. and Austin, S. (1991) Fine-structure analysis of the P1 plasmid partition site. *J. Bacteriol.*, **173**, 3630–3634.
20. Mori, H., Mori, Y., Ichinose, C., Niki, H., Ogura, T., Kato, A. and Hiraga, S. (1989) Purification and characterization of *SopA* and *SopB* proteins essential for F plasmid partitioning. *J. Biol. Chem.*, **264**, 15535–15541.
21. Schumacher, M.A. and Funnell, B.E. (2005) Structures of *ParB* bound to DNA reveal mechanism of partition complex formation. *Nature*, **438**, 516–519.
22. Khare, D., Ziegler, G., Lanka, E. and Heinemann, U. (2004) Sequence-specific DNA binding determined by contacts outside the helix-turn-helix motif of the *ParB* homolog *KorB*. *Nat. Struct. Mol. Biol.*, **11**, 656–663.
23. Ravin, N.V., Reeh, J. and Lane, D. (2003) Mapping of functional domains in F plasmid partition proteins reveals a bipartite *SopB*-recognition domain in *SopA*. *J. Mol. Biol.*, **329**, 875–889.
24. Hanai, R., Liu, R.P., Benedetti, P., Caron, P.R., Lynch, A.S. and Wang, J.C. (1996) Molecular dissection of a protein *SopB* essential for *Escherichia coli* F plasmid partition. *J. Biol. Chem.*, **271**, 17469–17475.
25. Lynch, A.S. and Wang, J.C. (1995) *SopB* protein-mediated silencing of genes linked to the *sopC* locus of *Escherichia coli* F plasmid. *Proc. Natl Acad. Sci. USA*, **92**, 1896–1900.
26. Schumacher, M.A., Piro, K.M. and Xu, W. (2010) Insight into F plasmid DNA segregation revealed by structures of *SopB* and *SopB*-DNA complexes. *Nucleic Acids Res.*, **38**, 4514–4526.
27. Pillet, F., Sanchez, A., Lane, D., Anton Leberre, V. and Bouet, J.Y. (2011) Centromere binding specificity in assembly of the F plasmid partition complex. *Nucleic Acids Res.*, **39**, 7477–7486.
28. Lemonnier, M., Bouet, J.Y., Libante, V. and Lane, D. (2000) Disruption of the F plasmid partition complex in vivo by partition protein *SopA*. *Mol. Microbiol.*, **38**, 493–505.
29. Godfrin-Estevan, A.M., Pasta, F. and Lane, D. (2002) The *parAB* gene products of *Pseudomonas putida* exhibit partition activity in both *P. putida* and *Escherichia coli*. *Mol. Microbiol.*, **43**, 39–49.
30. Dubarry, N., Pasta, F. and Lane, D. (2006) *PurABS* systems of the four replicons of *Burkholderia cenocepacia*: new chromosome centromeres confer partition specificity. *J. Bacteriol.*, **188**, 1489–1496.
31. Luscombe, N.M., Laskowski, R.A. and Thornton, J.M. (2001) Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. *Nucleic Acids Res.*, **29**, 2860–2874.
32. Gordon, G.S., Sitnikov, D., Webb, C.D., Teleman, A., Straight, A., Losick, R., Murray, A.W. and Wright, A. (1997) Chromosome and low copy plasmid segregation in *E. coli*: visual evidence for distinct mechanisms. *Cell*, **90**, 1113–1121.
33. Niki, H. and Hiraga, S. (1997) Subcellular distribution of actively partitioning F plasmid during the cell division cycle in *E. coli*. *Cell*, **90**, 951–957.
34. Leonard, T.A., Butler, P.J. and Lowe, J. (2004) Structural analysis of the chromosome segregation protein *SpoB* from *Thermus thermophilus*. *Mol. Microbiol.*, **53**, 419–432.
35. Leonard, T.A., Butler, P.J. and Lowe, J. (2005) Bacterial chromosome segregation: structure and DNA binding of the *Soj* dimer—a conserved biological switch. *EMBO J.*, **24**, 270–282.
36. Yamaichi, Y. and Niki, H. (2000) Active segregation by the *Bacillus subtilis* partitioning system in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **97**, 14656–14661.
37. Funnell, B.E. and Gagnier, L. (1993) The P1 plasmid partition complex at *parS*: II. Analysis of *ParB* protein binding activity and specificity. *J. Biol. Chem.*, **268**, 3616–3624.
38. Bouet, J.Y., Surtees, J.A. and Funnell, B.E. (2000) Stoichiometry of P1 plasmid partition complexes. *J. Biol. Chem.*, **275**, 8213–8219.
39. Passot, F.M., Calderon, V., Fichant, G., Lane, D. and Pasta, F. (2012) Centromere binding and evolution of chromosomal partition systems in the *Burkholderiales*. *J. Bacteriol.*, **194**, 3426–3436.
40. Golovanov, A.P., Barilla, D., Golovanova, M., Hayes, F. and Lian, L.Y. (2003) *ParG*, a protein required for active partition of bacterial plasmids, has a dimeric ribbon-helix-helix structure. *Mol. Microbiol.*, **50**, 1141–1153.
41. Woodecock, D.M., Crowther, P.J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S.S., Michael, M.Z. and Graham, M.W. (1989) Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Res.*, **17**, 3469–3478.
42. Grant, S.G., Jesse, J., Bloom, F.R. and Hanahan, D. (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl Acad. Sci. USA*, **87**, 4645–4649.
43. Bouet, J.Y., Bouvier, M. and Lane, D. (2006) Concerted action of plasmid maintenance functions: partition complexes create a requirement for dimer resolution. *Mol. Microbiol.*, **62**, 1447–1459.
44. Casadaban, M.J. and Cohen, S.N. (1980) Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J. Mol. Biol.*, **138**, 179–207.
45. Bouet, J.Y., Campo, N.J., Krusch, H.M. and Louarn, J.M. (1996) The effects on *Escherichia coli* of expression of the cloned bacteriophage T4 nucleoid disruption (*nud*) gene. *Mol. Microbiol.*, **20**, 519–528.
46. Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M. 2nd and Peterson, K.M. (1995) Four new derivatives of the broad-host-range cloning vector *pBBR1MCS*, carrying different antibiotic-resistance cassettes. *Gene*, **166**, 175–176.
47. Ravin, N. and Lane, D. (1999) Partition of the linear plasmid N15: interactions of N15 partition functions with the *sop* locus of the F plasmid. *J. Bacteriol.*, **181**, 6898–6906.
48. Lin, D.C.H. and Grossman, A.D. (1998) Identification and characterization of a bacterial chromosome partitioning site. *Cell*, **92**, 675–685.

Supplementary data

Insight into centromere binding properties of ParB proteins:
a secondary binding motif is essential for bacterial genome maintenance

Aurore Sanchez, Jérôme Rech, Cyrielle Gasc and Jean-Yves Bouet*

* For correspondence: Jean-Yves Bouet, LMGM, CNRS, Université Paul Sabatier, Bât. IBCG, 118 route de Narbonne, 31062 Toulouse Cedex 9, France.

Tel: (+33) 561 335 906; Fax: (+33) 561 335 886; E-mail: jean-yves.bouet@ibcg.biotoul.fr

Supplementary data content:

| | |
|---|------|
| Figure S1: <i>Prediction of the interaction between R219 and guanine 7</i> | p. 2 |
| Figure S2: <i>Stability assays of mFs carrying WT or R219 SopB alleles</i> | p. 3 |
| Figure S3: <i>Purification steps of wild-type and R219 SopB variants</i> | p. 4 |
| Figure S4: <i>Plasmid construction and maps</i> | p. 5 |
| Table SI: <i>E. coli</i> strains used in this study | p. 7 |
| Table SII: <i>Plasmids used in this study</i> | p. 8 |
| Supplementary references | P. 9 |

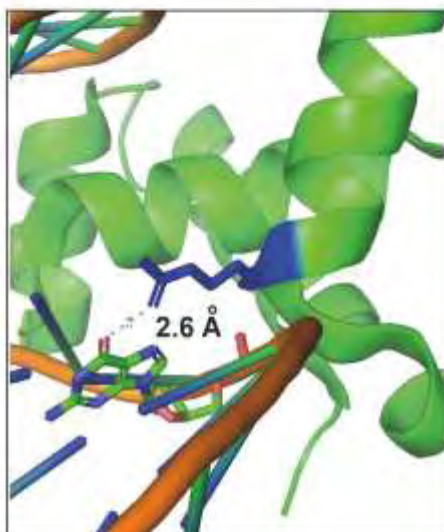


Figure S1: *Prediction of the interaction between R219 and guanine 7.* Details of the SopB-*sopC* structure, drawn with PyMOL from accession number 3EZf (26), are focused on the R219 residue (in blue). A distance of 2.6 angströms is estimated between the NH group of R219 and the O group of guanine. Notice that the resolution of the co-crystal structure was determined at ~3 angströms (26).

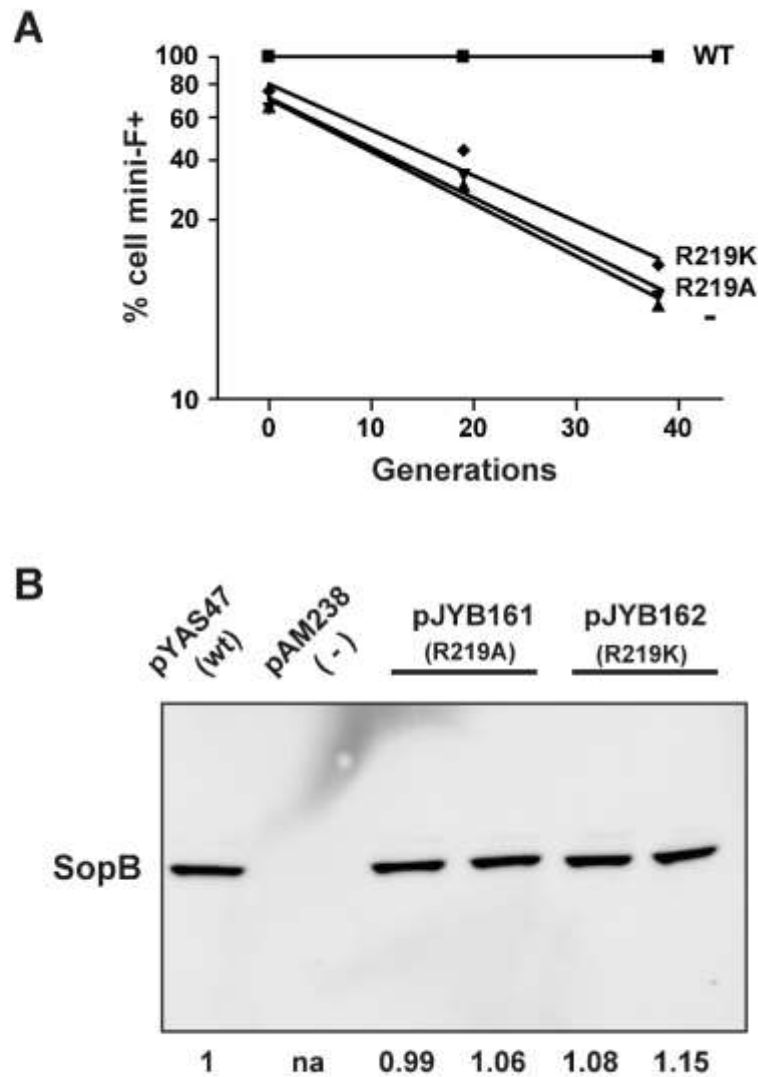


Figure S2: Stability assays of mFs carrying WT or R219 SopB alleles. **A-** The mini-F plasmid is not stabilized by *sopB*-R219 alleles. The percentage of cells retaining the mini-F (pDAG209) without selective pressure was assayed over 40 generations. SopA and SopB proteins were provided *in trans* from pSC101 derivatives: pYAS47 (*sopB*; square), pAM238 (-; triangle); pJYB161 (*sopB*-R219A; inverted triangle); pJYB162 (*sopB*-R129K; diamond). Fitted curves of representative experiments are drawn. **B-** SopB variants are expressed at the same level as WT SopB. Western blot analysis of SopB WT and variants R129A and R219K expressed from pYAS47, pJYB161 and pJYB162, respectively, were performed on samples taken from stability assay at time zero. pAM238 vector expressed no Sop proteins (-). Samples were resolved on a 4-12% Bis-Tris gradient gel (Invitrogen), and SopBs, revealed using anti-SopB antibodies, were quantified. Quantification of SopB levels relative to wild type SopB expressed from pYAS47 (1) was indicated below.

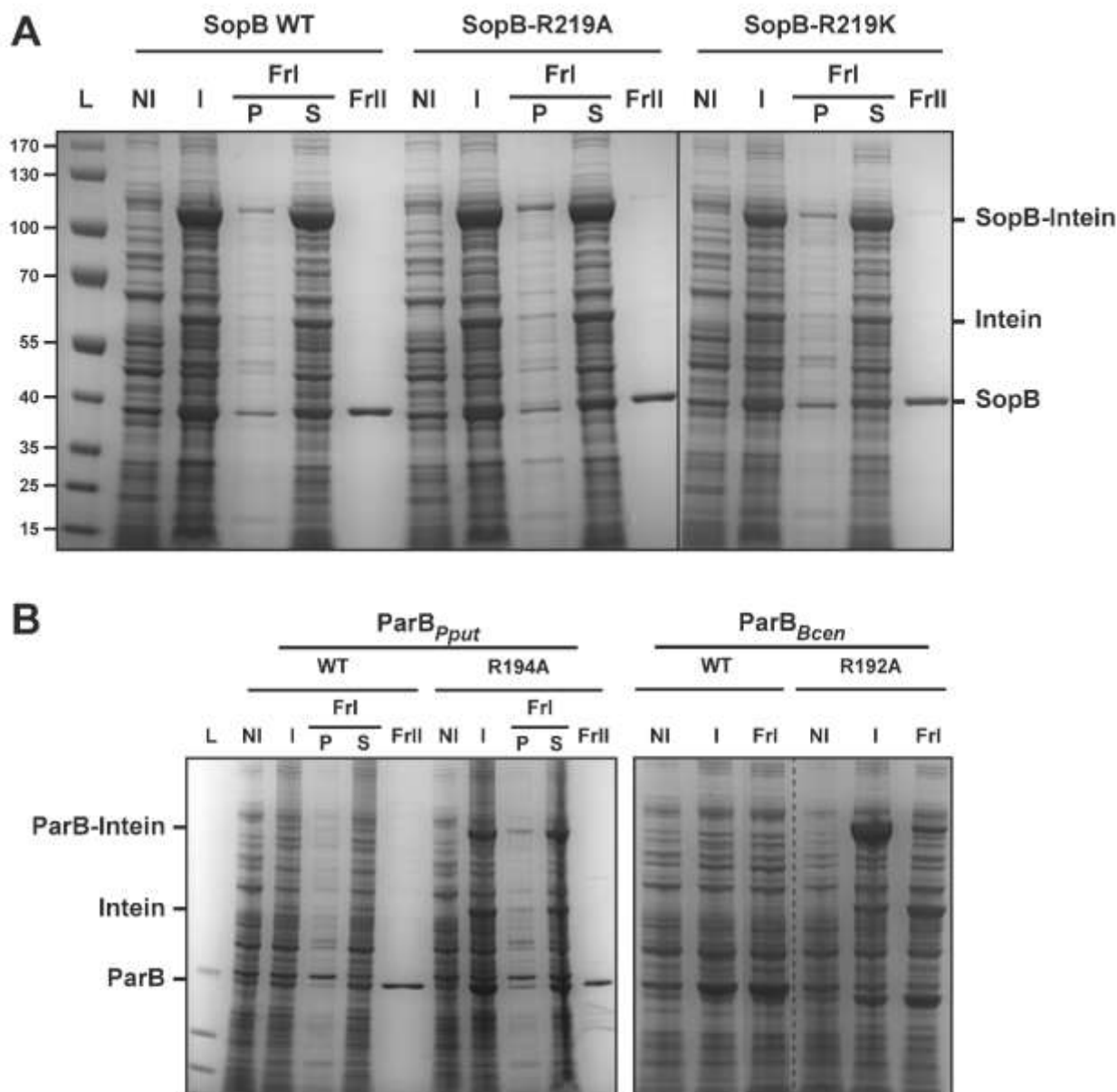
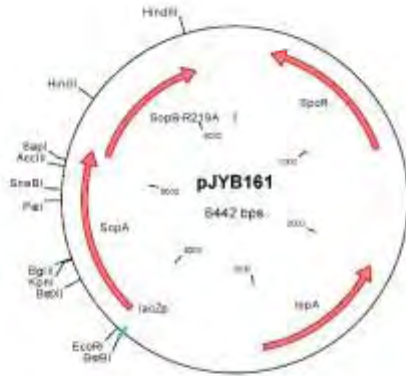


Figure S3: Purification steps of proteins used in this study. All the protein purifications were performed from Intein-tagged recombinant proteins. **A** - Purification of wild-type and R219 SopB variants. **B** - Purification of wild-type and R194A or R192A ParB variants from *Pput* or *Bcen*, respectively. Commassie-stained gels for each purified protein included non-induced (NI) and induced (I) cultures, supernatant (S) and pellet (P) of Fraction I (FrI) obtained after centrifugation of the crude lysate, and the fraction II (FrII) corresponding to the elution from the chitin column after the cleavage of the Intein tag. Because WT ParB_{Bcen} protein does not bind to the chitin column, FrI were directly incubated with DTT to induce the Intein cleavage. Molecular weights (Kda) of the ladder (L) proteins are indicated on the left.

Figure S4: Plasmid construction and maps.

- **pJYB161** (SopB-R219A) was constructed by Quick change mutagenesis from pYAS47 (Ah-Seng, unpublished) using the following oligonucleotides:
 SopB-R219Afl: 5'-CGGTGAACTATCTGCCgcaTCAGGT-3'
 SopB-R219Ar1: 5'-GAAGTGCATCACCTgaTGCGGCAGA-3'



- **pJYB162** (SopB-R219K) was constructed as pJYB161 using:
 SopB-R219Kfl: 5'-CGGTGAACTATCTGCCaaaTCAGGT-3'
 SopB-R219Kr1: 5'-GAAGTGCATCACCTGAttGGCAGA-3'

- **pJYB167** (ParB_{Bcen-C1}-R192A) was constructed by Quick change mutagenesis from pDAG562 using:
 AS32L: 5'-CATGGGGGCACGCagcaGCGCTGCTCGCGGTC-3'
 AS32R: 5'-ACCGCGAGCAGCGCtgcTGCGTGCCCCATG-3'

- **pJYB168** (ParB_{Pput}-R194A) was constructed by Quick change mutagenesis from pDAG503 using:
 AS33L: 5'-GAGATGGGGCACGCCgcaGCATTGCTGGGCCTG-3'
 AS33R: 5'-CAGGCCAGCAATGCTgcGGCGTGCCCCATCTC-3'

- **pAS13** (ParB_{Pput}) was constructed by insertion of a PCR fragment, amplified from pDAG503, in the pTYB1 digested by *NdeI* and *KpnI* restriction enzymes (sites indicated in bold) using:
 AS45L: 5'-AGAAGGAGATATAC**ATAT**GGCCGTCAGAAACGGGGTCTC-3'
 AS45R: 5'-CATTAAACATT**GGTACC**CTTGCAAAGCACCCGCGGATGTGAGCGAGCACGCC-3'

- **pAS14** (ParB-R194A_{Pput}) was constructed as pAS13 using pJYB168 as DNA template for PCR amplification.

- **pAS15** (ParB_{Bcen-C1}) was constructed as pDag562 by insertion of a PCR fragment, amplified from pJYB562, using:
 AS46L: 5'-AGAAGGAGATATAC**ATAT**GAGCGAAGCGAGGAAGAAAGAC-3'
 AS46R: 5'-CATTAAACATT**GGTACC**CTTGCAAAGCACCCCTCTTCGGTCGCGACGTTGCC-3'

- **pAS16** (ParB-R192A_{Bccn}) was constructed as pAS15 using pJYB167 as DNA template for PCR amplification.

- **pJYB208** (SopB-R219K) was constructed by Quick change mutagenesis from pYAS6 using the following oligonucleotides:

sopB219Afl: 5'-CGGTGAACTATCTGCCgcaTCAGGT-3'

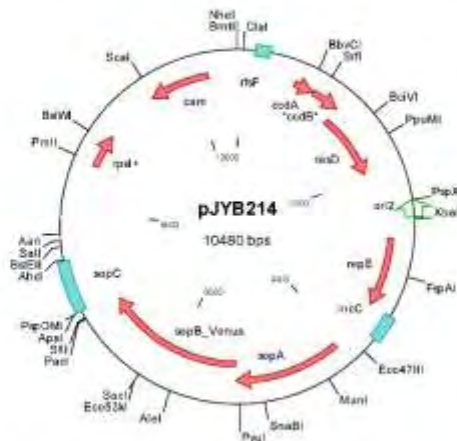
sopB219Arl: 5'-GAAGTGCATCACctgaTGCGGCAGA-3'

- **pJYB209** (SopB-R219K) was constructed as pJYB208 using:

sopB219Kfl: 5'-CGGTGAACTATCTGCCaaaTCAGGT-3'

sopB219Krl: 5'-GAAGTGCATCACCTGAtttGGCAGA-3'

- **pJYB214** was constructed using the In-Fusion HD Cloning System (Clontech) by insertion of two PCR fragments in the pDAG114 mini-F digested by *PvuI* and *ApaI* restriction enzymes (cloning information available upon request). Note that the CcdB toxin is inactive due to directed mutagenesis in the parental pDAG114 (47).



- **pJYB223** was constructed by allelic replacement of SopB from pJYB161 using *SacI* and *PvuI* in pJYB214 digested with the same restrictions.

Table S1: *E. coli* strains used in this study.

| Strains | Genotype/relevant properties | Source/references |
|--------------|---|-------------------|
| C2833 | <i>F'</i> , <i>proA</i> ⁺ <i>B</i> ⁺ , <i>lacF</i> ^{zzf::Tn10(Tet^R)/<i>fhuA2</i>, <i>ompT</i>, <i>lon</i>, <i>dcm</i>, <i>gal</i>, <i>lacZ</i>::<i>gene1</i> T7, <i>suLA11</i>, <i>endA1</i>, Δ(<i>mcr-73::mTn10,Tc^S</i>), Δ(<i>zgb-210::Tn10, Tc^S</i>), Δ(<i>mcrC-mrr</i>)114::IS10} | NEBioLabs |
| DH5 α | <i>F</i> , <i>endA1</i> , <i>hsdR17</i> , <i>supE44</i> , <i>thi-1</i> , <i>recA1</i> , <i>gyrA</i> , <i>relA1</i> , <i>deoR</i> , Δ (<i>lacZYA-argF</i>)U169, <i>f80dLac</i> Δ (<i>lacZ</i>)M15 | (41) |
| DH10B | <i>F</i> , <i>araD139</i> , Δ (<i>ara-leu</i>)7679, Δ <i>lacX74</i> , <i>galU</i> , <i>galK</i> , <i>rpsL</i> , <i>deoR</i> , <i>endA1</i> , <i>nupG</i> , <i>recA1</i> , <i>mcrA</i> , Δ (<i>mrr-hsdRMS-mcrBC</i>) | (42) |
| DLT1215 | <i>F</i> , <i>thi</i> , <i>leu</i> , <i>thyA</i> , <i>deoB</i> , <i>supE</i> , Δ (<i>ara-leu</i>)7696, <i>zac3051::Tn10</i> , <i>rpsL812</i> | (43) |
| DLT2905 | MC1061/pDAG209 | This work |
| DLT2906 | DLT2905/pAM238 | This work |
| DLT2907 | DLT2905/pYAS47 | This work |
| DLT2959 | DLT2905/pJYB161 | This work |
| DLT2960 | DLT2905/pJYB162 | This work |
| DLT2910 | C2833/pJYB208 | This work |
| DLT2911 | C2833/pJYB209 | This work |
| DLT2950 | DLT1215/pJYB214 | This work |
| DLT2951 | DLT1215/pJYB223 | This work |
| DLT3013 | C2833/pAS13 | This work |
| DLT3014 | C2833/pAS14 | This work |
| DLT3015 | C2833/pAS15 | This work |
| DLT3016 | C2833/pAS16 | This work |
| MC1061 | <i>araD139</i> , Δ (<i>ara-leu</i>)7679, Δ <i>lacX74</i> , <i>galU</i> , <i>galK</i> , <i>rpsL</i> , <i>thi</i> , <i>hsdR2</i> , <i>mcrB</i> | (44) |

Table SII: Plasmids used in this study.

| Plasmids | Relevant characteristics | Source/references |
|-----------|---|-------------------|
| pAM238 | <i>lacOPZ'</i> Ωmes in <i>EcoRI-HindIII</i> interval of pGB2, <i>aadA</i> ⁺ | (45) |
| pBBR1MCS5 | <i>rep_{pBBR1}</i> , <i>mobRK2</i> , <i>sacRB</i> , <i>strAB</i> | (46) |
| pAS13 | pTYB1, pT7:: <i>ParB_{Ypur}</i> -G291::See Intein::CBD | This work |
| pAS14 | pTYB1, pT7:: <i>ParB_{Ypur}</i> -G291-R194A::See Intein::CBD | This work |
| pAS15 | pTYB1, pT7:: <i>ParB_{Bcen}</i> -G306::See Intein::CBD | This work |
| pAS16 | pTYB1, pT7:: <i>ParB_{Fpur}</i> -G306-R192A::See Intein::CBD | This work |
| pDAG114 | <i>repFIA</i> ⁺ , <i>ccdB</i> ⁻ , <i>resD</i> ⁺ , <i>rsfF</i> ⁺ , <i>cat</i> ⁺ | (47) |
| pDAG209 | pDAG114 Δ <i>sopAB</i> | (43) |
| pDAG503 | pAM238 with <i>parAB</i> from <i>P. putida</i> | (29) |
| pDAG504 | pDAG114 Δ <i>sopABC</i> <i>parS_{Yp}</i> ⁺ | (29) |
| pDAG551 | pDAG114 Δ <i>sopABC</i> <i>parS</i> | (30) |
| pDAG562 | pBBR1MCS5 with <i>parAB</i> from <i>B. cenocepacia</i> chr1 | (30) |
| pJYB161 | pYAS47, <i>sopB</i> -R219A | This work |
| pJYB162 | pYAS47, <i>sopB</i> -R219K | This work |
| pJYB167 | pDAG562, <i>parB</i> -R192A (<i>Bcen</i>) | This work |
| pJYB168 | pDAG503, <i>parB</i> -R194A (<i>Pput</i>) | This work |
| pJYB208 | pYAS6, pT7:: <i>sopB</i> -R219A-G324::See Intein::CBD | This work |
| pJYB209 | pYAS6, pT7:: <i>sopB</i> -R219K-G324::See Intein::CBD | This work |
| pJYB214 | pDAG114, <i>sopB</i> ::Venus-Yfp | This work |
| pJYB223 | pJY214, <i>sopB</i> -R219A::Venus-Yfp | This work |
| pYAS6 | pTYB1, pT7:: <i>sopB</i> -G324::See Intein::CBD | (13) |
| pYAS47 | pAM238, <i>plac</i> :: <i>sopAB</i> | Ah-Seng, unpubl. |

Supplementary references

13. Ah-Seng, Y., Lopez, F., Pasta, F., Lane, D. and Bouet, J.Y. (2009) Dual role of DNA in regulating ATP hydrolysis by the SopA partition protein. *J Biol Chem*, **284**, 30067-30075.
29. Godfrin-Estevenon, A.M., Pasta, F. and Lane, D. (2002) The parAB gene products of *Pseudomonas putida* exhibit partition activity in both *P. putida* and *Escherichia coli*. *Mol Microbiol*, **43**, 39-49.
30. Dubarry, N., Pasta, F. and Lane, D. (2006) ParABS systems of the four replicons of *Burkholderia cenocepacia*: new chromosome centromeres confer partition specificity. *J Bacteriol*, **188**, 1489-1496.
41. Woodcock, D.M., Crowther, P.J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S.S., Michael, M.Z. and Graham, M.W. (1989) Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Res*, **17**, 3469-3478.
42. Grant, S.G., Jessee, J., Bloom, F.R. and Hanahan, D. (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc Natl Acad Sci U S A*, **87**, 4645-4649.
43. Bouet, J.Y., Bouvier, M. and Lane, D. (2006) Concerted action of plasmid maintenance functions: partition complexes create a requirement for dimer resolution. *Mol Microbiol*, **62**, 1447-1459.
44. Casadaban, M.J. and Cohen, S.N. (1980) Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J Mol Biol*, **138**, 179-207.
45. Bouet, J.Y., Campo, N.J., Krisch, H.M. and Louarn, J.M. (1996) The effects on *Escherichia coli* of expression of the cloned bacteriophage T4 nucleoid disruption (*ndd*) gene. *Mol Microbiol*, **20**, 519-528.
46. Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M., 2nd and Peterson, K.M. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene*, **166**, 175-176.
47. Ravin, N. and Lane, D. (1999) Partition of the linear plasmid N15: interactions of N15 partition functions with the sop locus of the F plasmid. *J Bacteriol*, **181**, 6898-6906.

Chapitre III

Organisation du complexe de partition "étendu"

I. Etat de l'Art

La première étape de la partition implique une fixation des protéines SopB à leur site centromérique, ce qui permet l'assemblage du "complexe de partition". Il apparaît aussi que la fixation des dimères de SopB sur les 10 sites de fixations *sopC* conduise à la formation d'un complexe de partition d'ordre supérieur que nous appellerons le complexe de partition "étendu". L'idée d'un tel complexe provient de diverses études génétiques, biochimiques et topologiques (cf. **Introduction, Partie III.D**).

SopB est présent avec un excès d'un facteur 15 à 20 par rapport au nombre de sites de liaison potentiels *sopC* dans la cellule (Bouet *et al.*, 2005). La question de la localisation des SopB excédentaires est naturellement apparue. *In vivo*, l'interaction de SopB avec un plasmide portant *sopC* induit un déficit d'environ douze tours négatifs (Lynch and Wang, 1994; Biek and Shi, 1994; Biek and Strings, 1995). Cette contrainte topologique existe même en présence d'un seul site de fixation *sopC*, suggérant l'existence d'un mécanisme de compensation expliquant le faible effet du nombre de sites de fixation sur cette contrainte topologique (Biek and Shi, 1994; Biek and Strings, 1995). Par ailleurs, SopB est capable d'éteindre des gènes adjacents à la séquence centromérique *sopC*, par un phénomène qui a été appelé "Silencing" (Lynch and Wang, 1995; Kim and Wang, 1999). L'explication de ce phénomène a été attribuée à la capacité des dimères de SopB à s'oligomériser entre eux puis s'étaler sur l'ADN plasmidique à partir du site centromérique, *sopC*.

Le phénomène de "Silencing", l'excès de protéine SopB dans la cellule et les contraintes topologiques que la fixation de SopB impose, sont autant d'arguments indiquant que SopB pourrait être recruté au niveau du centromère pour former un complexe "étendu". L'organisation d'un tel complexe n'est à ce jour connu pour aucun des systèmes de partition de type I. Il a d'abord été proposé que la fixation des SopB sur le site centromérique et sa

propagation sur l'ADN flanquant, permette un enroulement de l'ADN autour des protéines SopB. Un mécanisme similaire a été décrit pour les nucléosomes eucaryotes. Cependant, cette hypothèse a été réfutée par des travaux réalisés au sein de notre équipe (Bouet & Lane 2009) (cf. Introduction, Partie III.D.1).

L'étude présentée dans ce chapitre vise à comprendre les modalités d'étalement de SopB et si cet étalement suit une organisation précise pouvant être impliquée dans le processus de ségrégation.

Une expérience de retard de migration sur gel décrite dans la littérature, montre la capacité de SopB à se fixer spécifiquement à son centromère mais révèle également la capacité de SopB à former des retards de plus haut poids moléculaires (Ah-Seng *et al.*, 2009). Ces complexes pourraient correspondre à de l'étalement de SopB, résultant de la polymérisation de SopB sur l'ADN à la manière d'un protéo-nucléofilament qui s'initierait au niveau du centromère.

En se basant sur cette étude et par une approche *in vitro*, nous avons caractérisé ces complexes de haut poids moléculaires. Cet étalement nécessite deux types d'interaction, des interactions ADN/protéine et protéine/protéine. L'impossibilité à obtenir des empreintes visualisables de la protéine SopB sur l'ADN flanquant la séquence centromérique, semble indiquer que cet étalement n'est pas régulier à la manière d'un proté-nucléofilament comme le présentait l'ancien modèle. Par une approche *in vivo* d'immunoprécipitation de chromatine (ChIP), couplée à des techniques de hautes détections (qChIP et ChIP-seq), nous avons montré de manière directe que SopB est capable de s'étendre sur une distance de 20 kilobases de part et d'autre du centromère. En collaboration avec l'équipe du Dr Marcelo Nollmann au Centre de Biochimie Structurale de Montpellier pour l'utilisation de la microscopie de super-résolution (PALM), nous avons montré que les protéines SopB sont fortement concentrées dans des "clusters" de 40nm ce qui suggère un niveau d'organisation élevé.

L'ensemble de ces travaux, nous ont permis de proposer un nouveau modèle dans lequel le complexe de partition "étendu" serait une structure concentrée, dynamique et résultant d'interaction stochastique entre SopB et l'ADN avoisinant *sopC*.

II. Présentation de l'article : "The bacterial segrosome assembles without requiring ParB filament formation"

Authors: Aurore Sanchez, Diego I. Cattoni, Jérôme Rech, Marcelo Nollmann and Jean-Yves Bouet

Keywords: DNA segregation, Plasmid partition, Centromere, ParB, superresolution microscopy, PALM, ParB nucleofilament, spreading, bacteria, *E. coli*, partition complex

A. Abstract

In bacteria, assembly of a nucleoprotein complex on centromere DNA sequences is the first step in the active partition process that ensures a bidirectional segregation of DNA molecules. Centromere binding proteins, ParB, display the capability to propagate along the DNA from their nucleation at the centromere. However, the organization and architecture of such assembly is not known for any type I partition system, the most widespread type of prokaryotic partition system. Here, we investigated the overall organization of the protein assembly on the centromere and adjacent DNA by combining different approaches including high-throughput sequencing, super-resolution microscopy and biochemical analyses, and we determined the mode of propagation of SopB nucleated at the centromere. Our results are inconsistent with the formation of a ParB filament nucleated at the centromere. Rather, our analyses of the plasmid F partition system reveal that the partition complex results from the stochastic binding of the centromere binding protein on neighboring *sopC* DNA.

B. Introduction

Chromosome segregation is a critical step for cell cycle progression in all cell types. In eukaryotes, mitosis ensures this activity into well-defined stages after chromosomes replication. By contrast, in bacteria this activity is not precisely separated into defined cell-cycle steps, although chromosomes are segregated in a very organized process concomitant with replication (for review, Kusminov, 2013). Active DNA segregation in most bacteria is driven by a three components system, generically called ParABS: the ParB family members recognize specific centromere-like DNA sequences (*parS*) and assemble a partition complex leading to the formation of a higher-order nucleoprotein complex (hereafter extended partition complex). The extended partition complex enables the segregation of plasmids or chromosomes to

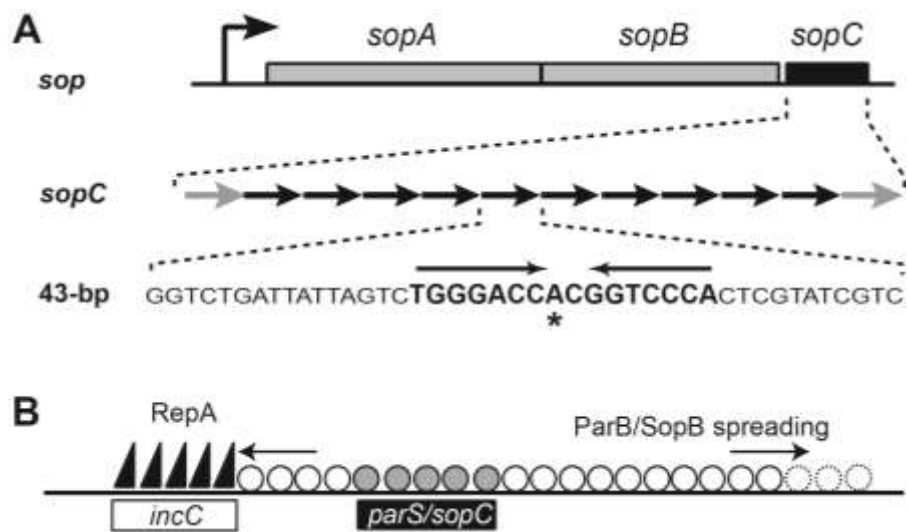


Figure 34 | Schematic representation of the *F* plasmid *sop* operon and current *ParB* spreading model.

(A) The *sop* operon containing the *sopA* and *sopB* genes, the *cis*-acting *sopC* site, and the promoter (arrow) region are drawn (not to scale). The *sopC* centromere is composed of 12 direct repeats of 43-bp (head to tail arrows). Each repeat contains a 16-bp *SopB* binding site (in bold) composed of two inverted repeat (head to head arrows) separated by a non-palindromic AC central. Note that the first and last repeats (gray arrows) are too divergent for binding to *SopB* *in vitro* (Pillet et al. 2011). (B) A protein-DNA complex acting as a roadblock to protein spreading along the DNA. In the schematic diagram, gray and open circles represent *ParB*-bound to *parS* and non-specific DNA, respectively; Black triangles represent *RepA* protein tightly bound to *incC* iteron site preventing *ParB* spreading to extend beyond this assembly. Model adapted from Rodionov et al. 1999.

daughter cells by interacting with a cognate protein, ParA (reviewed in Gerdes *et al.* 2010). The mechanisms by which ParB self-assembles into higher order complexes necessary for efficient chromosome segregation, as well as the architecture of such complexes remain elusive.

In bacteria, active segregation relies on either one of the three known types of DNA segregation systems, types **I** to **III**, according to the NTPase family to which ParA belongs to, Walker A cytoskeletal ATPase (ParA), actin-like ATPases (ParM), and tubulin-like GTPases (TubZ), respectively. Type **I** partition system is the most prevalent type on plasmids and the only type present on bacterial chromosomes. Its mechanism of action is still not established and the two main models – "filament pulling" or "diffusion-ratchet" based mechanisms - has been reviewed recently (Vecchiarelli *et al.* 2012). Interactions between ParA and the extended partition complex are essential and thus determining the global architecture of this complex is necessary to decipher the active partition mechanism.

After replication, the first step of the partition mechanism is the formation of the extended partition complex, that will serve not only to link centromere DNA to ParAs (Bouet and Funnell, 1999) but also to activate the dynamic properties of ParA that will provide the force necessary to segregate the plasmid copies (reviewed in Gerdes *et al.* 2010). While the primary binding of ParB on centromere DNAs is well characterized (Bouet *et al.* 2000; Funnell, 1991; Lin and Grossman, 1998; Pillet *et al.* 2011; Schumacher and Funnell, 2005; Schumacher *et al.* 2007b), the overall organization and architecture of partition complexes is not known for any type **I** partition system. This is in striking contrast to types **II** and **III** for which partition complexes architectures are well defined structurally, and appears as DNA-protein filaments with the centromeres wrapped around a helical scaffold of the ParBs (Aylett and Lowe, 2012; Moller-Jensen *et al.* 2007; Salje and Lowe, 2008; Schumacher *et al.* 2007a). The possibility that type **I** partition complexes could also assemble in nucleoprotein filaments organized in a solenoid structure was suggested by the capability of type **I** ParB to propagate from centromere sequences into flanking DNA (Biek and Shi, 1994; Rodionov *et al.* 1999).

The archetypical type **I** partition locus of the conjugative plasmid F of *Escherichia coli* is composed of the ParA homolog, SopA, the centromere binding protein, SopB, and the centromere site, *sopC*, composed of twelve direct repeats regularly spaced every 43-bp (**Figure 34A**). SopB dimers bind specifically with high affinity to 16-bp inverted repeats present in each direct repeat. The capability of SopB to propagate outside *sopC* was first detected by its a

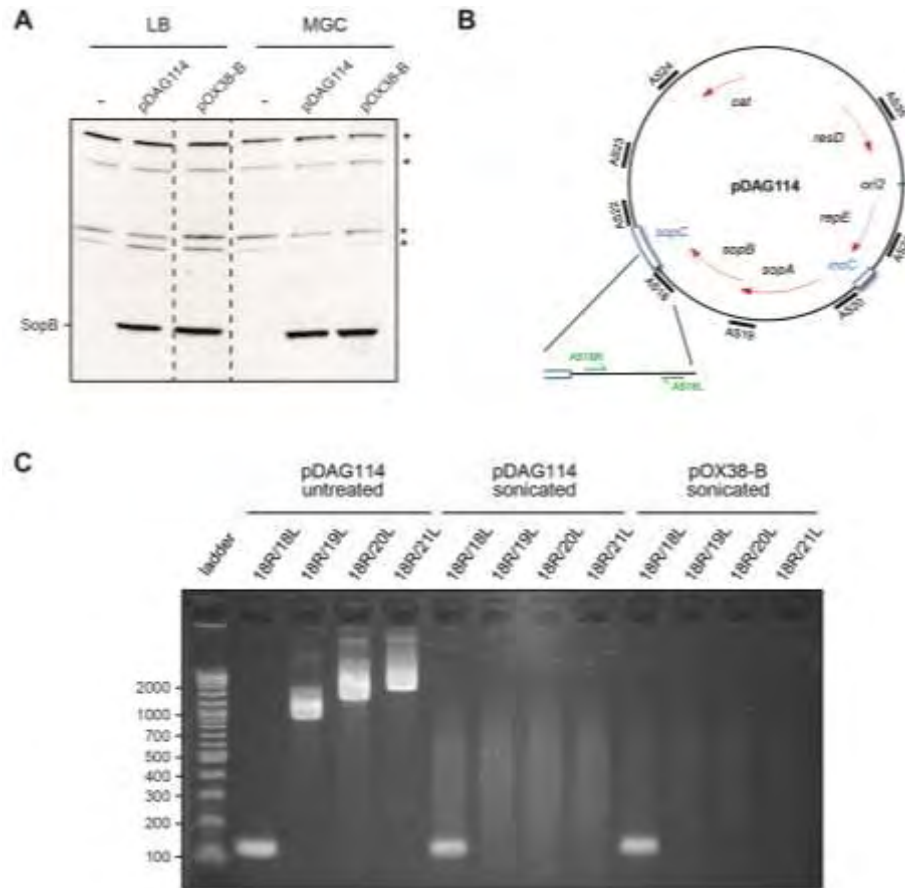


Figure 35 | Intracellular SopB proteins level and control of the sonication step for ChIP assays.

(A) Intracellular SopB level is similar when produced from the mini-F or the F plasmids. Cell extracts from strain DLT1215 carrying the mini-F (pDAG114), the F-plasmid (pOX38-B) or without plasmid (-) were fractionated by SDS-PAGE and subjected to Western blot analysis with anti-SopB antibody. Strains were grown in LB or in M9 glucose CSA (MGC). Stars (*) on the right indicate non-specific cross reactive bands which are indicative of the relative protein amount in each lane. (B) Schematic representation of mini-F and positioning of the primer pairs used in q-ChIP assays. Mini-F plasmid (pDAG114) is represented as a circle with genes and DNA sites depicted with red arrows and blue boxes, respectively. PCR-amplified DNA zones are represented with black lines and named according to the primer pairs used. For example, primer pairs AS18R and AS18L (green arrows) amplify the AS18 zone, as illustrated in the zoom near *sopC*. (C) Extensive sonication prevents PCR amplification between two adjacent primer pairs. PCR were performed on DNA from formaldehyde-treated cells prior (untreated) or after (sonicated) sonication of total DNA from cells carrying the mini-F (pDAG114) or plasmid F (pOX38-B). Different combinations of left (L) and right (R) primers were used. By contrast to the untreated sample, the absence of PCR products longer than 1 Kb in the sonicated samples indicated that the sonication step was efficient to provide a homogenous population of small DNA fragments.

marked effect on plasmid DNA supercoiling (Biek and Shi, 1994) and by its gene silencing effect in the vicinity of the F centromere, *sopC* (Lynch and Wang, 1995). Subsequently, in the case of plasmid P1, cross-linking experiments demonstrated that ParB-mediated silencing involves a direct interaction between ParB proteins and the silenced region (Rodionov et al. 1999). Similar results with chromosomal Par systems were also observed (Kusiak et al. 2011; Murray et al. 2006). At physiological levels, ParB was found to cover large region flanking *parS* sites on plasmid (up to 12 Kb; Rodionov et al. 1999) and on chromosome (> 10 Kb; Breier and Grossman, 2007). Interestingly, ParB propagation was shown to be blocked by protein-DNA complexes (Murray et al. 2006; Rodionov et al. 1999), leading to the hypothesis that ParB spreads laterally as a filament onto DNA adjacent to the centromere, as depicted in **Figure 34B**. Attempts to observe type I partition complexes by EM or AFM have not yet revealed their ultra-structural architecture (Hanai et al. 1996; Pratto et al. 2009; Wu et al. 2011).

In the present study, we investigated the overall organization of the SopB-*sopC* extended partition complex of plasmid F. Combining different approaches, we characterized the mode of propagation of SopB nucleated at the centromere. Superresolution microscopy revealed for the first time at the single molecule level the intracellular localization of SopB molecules. Collectively, our data allow defining the architecture of the partition complex and we propose a model in which the extended partition complex results from the stochastic binding of SopB on the DNA in the vicinity of the centromere.

C. Results

1. SopB propagates over large region on both F and mini-F plasmids

To investigate how much DNA flanking the plasmid F centromere is involved in the extended partition complex, we performed formaldehyde-mediated DNA-protein crosslinking followed by immunoprecipitation (ChIP) of SopB complexes, using strain DLT1215 carrying either a mini-F plasmid (pDAG114) or a F plasmid (pOX38-B). SopB is expressed at the same level from both plasmids (**Figure 35**). After formaldehyde treatment, exponentially growing cells were sonicated to shear DNA to ~400-bp. To ensure the observed enrichment pattern was not due to incomplete shearing, we controlled that primers from adjacent pairs could not amplify the sonicated but the untreated DNA (**Figure 35C**). DNA-bound SopB was immunoprecipitated with anti-SopB antibodies and the isolated DNA was then analyzed by

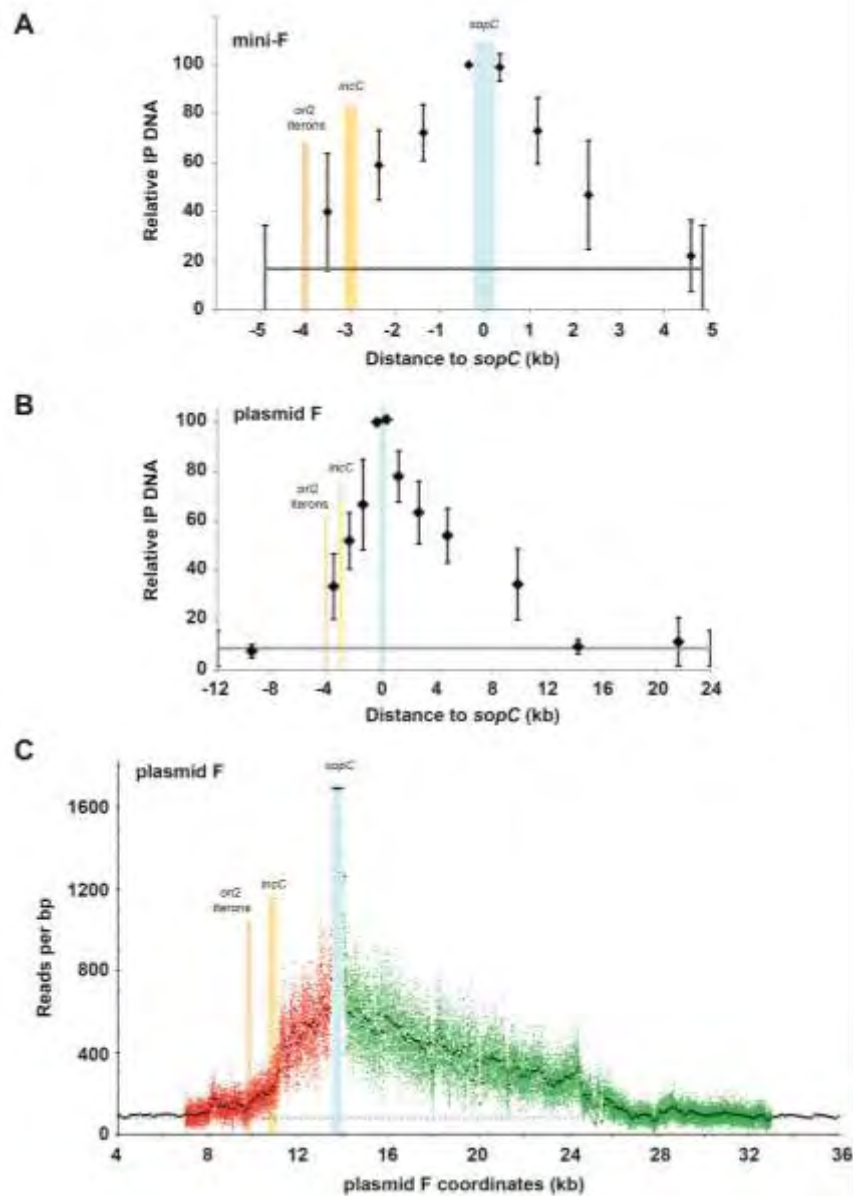


Figure 36 *SopB binds DNA outside *sopC* asymmetrically.*

Formaldehyde-treated cells carrying the mini-F or the plasmid F were immunoprecipitated with anti-SopB antibodies. The DNA recovered was quantified by q-PCR (A and B; data are average of three independent experiments) or by deep sequencing (C). The blue, yellow and rose vertical boxes represent *sopC*, *ori2* iterons and *incC* positioning, respectively.

(A) SopB spreading on the mini-F. The relative amount of DNA immunoprecipitated (IP DNA) quantified by q-PCR is plotted versus the nucleotide distance in base pair (Kb) to the center of *sopC*. The primer pairs are represented as black diamonds, and the most proximal to the left side of *sopC* (AS18) was taken as reference. The black diamond is the duplication of the primer pair diametrically opposite to *sopC* (coordinate 4.8 Kb). The baseline level (horizontal grey line) is determined by the primer pair *terC2* that amplifies DNA from the *E. coli* chromosome. (B) SopB spreading on the plasmid F measured as in (A). (C) SopB spreading on the plasmid F measured by ChIP-sequencing. The number of reads is plotted versus the nucleotide coordinates of pOX38-B. For clarity, only the reads between 7 to 13.7 Kb (red dots) and 14.8 to 33.3 Kb (green dots) are shown. The black dots correspond to the average number of reads per 100-bp windows. The baseline level (90 reads/bp; dashed grey line) was determined by averaging reads/bp on the plasmid F with the exclusion of the SopB binding zone (7 to 27 Kb). Note that due to sequence repetition (Figure 34A), reads carrying *sopC* sequences could not be directly positioned onto the F sequence. Therefore, the total number of reads that matches to *sopC* repeats were divided by the number of nucleotides, leading to the average of 1692 reads/bp over the centromere sequence.

quantitative polymerase chain reaction (q-PCR) utilizing several primer pairs that surround the *sopC* site (**Figure 35B and Table SI**).

On the mini-F, the DNA adjacent on both sides of *sopC* was efficiently immunoprecipitated (**Figure 36A**), indicating that SopB efficiently propagates outside the centromere. We observed that the level of IP DNA decreases as the distance from *sopC* increases. The primer pair diametrically positioned relative to *sopC* (4.6 Kb) was slightly above the background level indicating that SopB binds over the 9.76 Kb mini-F. We obtained essentially the same results with the plasmid F (**Figure 36B**), which expresses SopB at the same level as the mini-F (**Figure 35**), but SopB is detected over a longer distance, between 10 Kb to 15 Kb, on the right side. SopB propagation along the DNA is thus asymmetric.

Propagation of ParB outside the centromere was originally thought to result from a ParB polymerization termed “spreading”, i.e. the auto-assembly of ParB proteins side by side along the DNA, nucleated at the centromere that is blocked by nucleoprotein complexes (Rodionov *et al.* 1999). Such organization for SopB binding on the DNA would imply that the amount of IP DNA from sequences just next to the centromere DNA would be at about the same level as inside it. Because we could not test it directly by q-PCR due to (i) the repeated composition of *sopC*, and (ii) its lack of resolution, we turned to ChIP-sequencing analyses which provide results at high resolution.

2. SopB do not propagate linearly along the DNA outside the centromere

The propagation mode of SopB nucleated at the centromere was assayed by ChIP-seq (**Figure 36C**). After formaldehyde treatments of F carrying cells, extensive sonication provides DNA fragments of 160-bp in average (**see Figure 37**) ensuring that no bias is due to long DNA fragments. Reads from high-throughput sequencing were aligned with the genome of reference and SopB-specific IP DNA was only found on the plasmid F (not shown). The background level over the F plasmid is ~90 reads/bp and the maximum of IP DNA culminates at *sopC* (1700 reads/bp), indicating that the signal/background ratio is high (~20). We observed the same pattern as above with an asymmetric distribution of SopB binding sites over ~5 Kb and ~15 Kb on the left and right sides of *sopC*, respectively. Strikingly, the transition between *sopC* and the adjacent sequences is very abrupt with the number of reads/bp decreasing from ~1700 to ~750 over only 20-bp (**Figure 36C**). After these rapid drops, the number of reads/bp

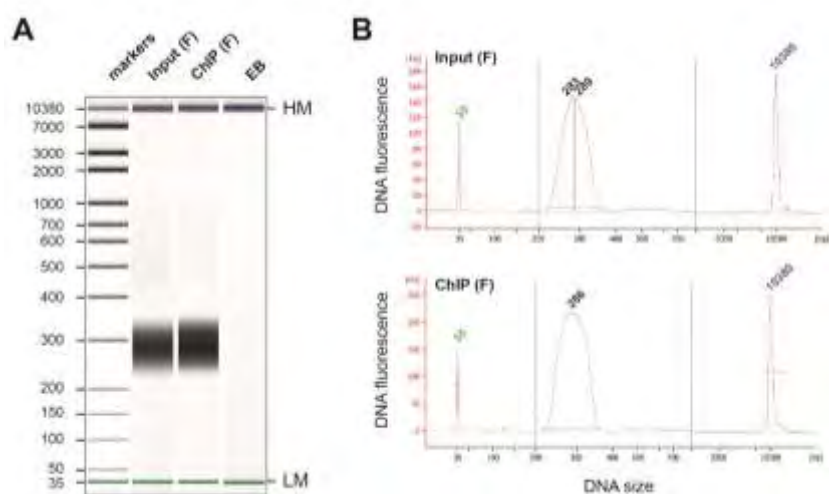


Figure 37 DNA profiles from formaldehyde-treated cells show extensive sonication prior to ChIP-sequencing analysis.

Each DNA library, prepared from the sonicated DNA fragments were ligated to DNA adapters (130-bp; Illumina), was analyzed by agarose gel electrophoresis and densitometric DNA profiles were determined (B). Low- and high-DNA markers, noted LM (35-bp) and HM (10380-bp), respectively, were added in each sample prior to electrophoresis. The average size of both samples (after subtraction of the adapter) was ~160-bp, ranging from ~100- to 250-bp.

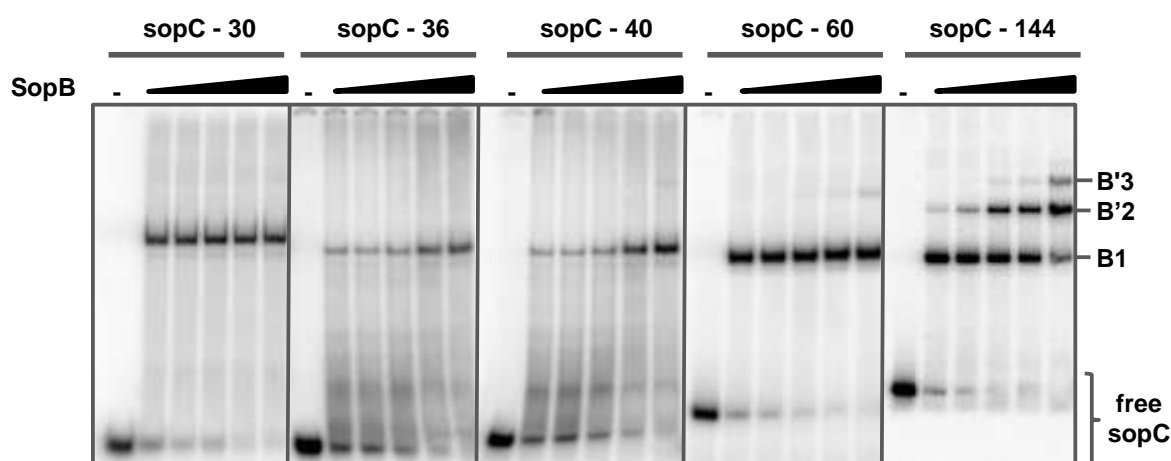


Figure 38 Formation of secondary SopB-DNA complexes requires long DNA fragments.

EMSA were performed with *sopC*-containing DNA fragments of increasing length (see Table SII). 32P-labeled DNA fragments carrying a single 16-bp SopB binding site, indicated on top of each gel, were incubated in the absence (-) or the presence of increasing concentrations of SopB (black triangle; 0, 10, 30, 100, 300, 1000 nM). All reaction mixtures, containing 100 µg.ml⁻¹ sonicated salmon sperm DNA, were analyzed by electrophoresis in a 6% polyacrylamide gel. The position of free and SopB-bound *sopC* probes are indicated on the right. B1 represents complexes involving the specific interaction on the 16-bp binding site while B'2 and B'3 complexes represent secondary complexes involving a specific site with one or two additional non-specific DNA binding interactions, respectively.

decreases progressively over large distance. This pattern in ChIP-reads on both sides of *sopC* suggests an organization of the extended partition complex without the formation of SopB nucleofilaments nucleated from the centromere.

3. RepE/iteron nucleoprotein complexes limit SopB spreading without acting as a barrier

ParB spreading as a nucleofilament was initially suggested by the finding that RepA/iterons complexes on plasmid P1 was shown to act as a roadblock for ParBP1 propagation (Rodionov et al. 1999). We thus looked at the effect of similar complexes on the pattern of SopB binding. At 2.5 and 3.5 Kb on the left side of *sopC* two arrays of binding sites for the replication initiator RepE, *incC* and *ori2* iterons (Uga et al. 1999), are present (represented as vertical orange bars in **Figure 36**), with the same organization and configuration relative to the centromere as for the homologous RepA/*inc* complexes of plasmid P1. On the plasmid F, SopB binds outside the iteron regions (**Figure 36B, 2C**) indicating that SopB propagation from the centromere is not stopped by the formation of RepE-*incC* complexes.

On the mini-F, the absence of RepE/iteron effect (**Figure 36A**) is probably due to the possibility that SopB binding at sites diametrically opposed to *sopC* could arise from both sides as expected from a plasmid whose size is lower than the 10-16 Kb propagation capacity observed with the larger plasmid F.

4. Formation of the extended partition complex requires both protein-protein and protein-DNA interactions

To further investigate the organization of the extended partition complex, we analyzed the formation of "secondary binding complexes" on centromere DNA containing only one SopB binding site using electromobility shift assay (EMSA; **Figure 38A**, left panel). As previously described, SopB specifically binds the 16-bp *sopC* with high affinity (2.5 nM) (Ah-Seng et al. 2009; Pillet et al. 2011). At 10 nM SopB, nearly all the *sopC* DNA probe was shifted in a discrete and specific complex (B1). A second discrete complex (B'2) was detected above 30 nM SopB and represented half of the shifted complexes at ~500 nM. A third complex (B'3) was also detected above 300 nM. We focused our investigation on the interaction required for the formation of the B'2 complex. Since EMSA were performed in the presence of large excess of non-specific DNA, the B'2 complex was not expected to arise from non-specific binding to the *sopC* probe. This was also excluded by (i) comparing the shifted pattern of a *sopC* and a non-

Figure 39 *SopB binding with specific and non-specific DNA probes.*

Mixtures were incubated at 30°C for 15 min in binding buffer containing 50 mM Hepes-KOH pH 7.5, 100 mM KCl, 100 µg/ml BSA and 1 mM DTT, supplemented (A) or not (B) with 100 µg/ml sonicated salmon sperm DNA, and then loaded on a 5 % polyacrylamide gel in Tris-Borate. Following electrophoresis, gels were dried and exposed to a phosphor screen for imaging with a Typhoon (GE Healthcare).

(A) Radiolabeled *sopC*-144 and NS-114, corresponding to a DNA fragment of 144-bp containing one SopB binding site and of 114-bp of non-specific DNA (repE), respectively, were incubated with increasing concentrations of SopB as indicated by the black triangle (10, 20, 40, 80, 160, 320, 640, 1280, 2560 and 5120 nM). The positions of free and SopB-bound *sopC*-144 probes are indicated on the left (B1, B'2, B'3 and B'4). The positions of free and SopB-bound NS-114 probes are indicated on the right (NS1 and NS2). Asterix (*) represents a contaminating band arising from the specific *sopC*-144 probes that migrates slightly above NS1.

(B) Radiolabeled NS-114 DNA fragment was incubated with increasing SopB concentrations (100, 200, 300, 400, 500, 600, 700, 800, 900, 1000 and 2000 nM). The positions of free and SopB-bound NS-114 probes are indicated on the left. HMW stand for high molecular weight complexes.

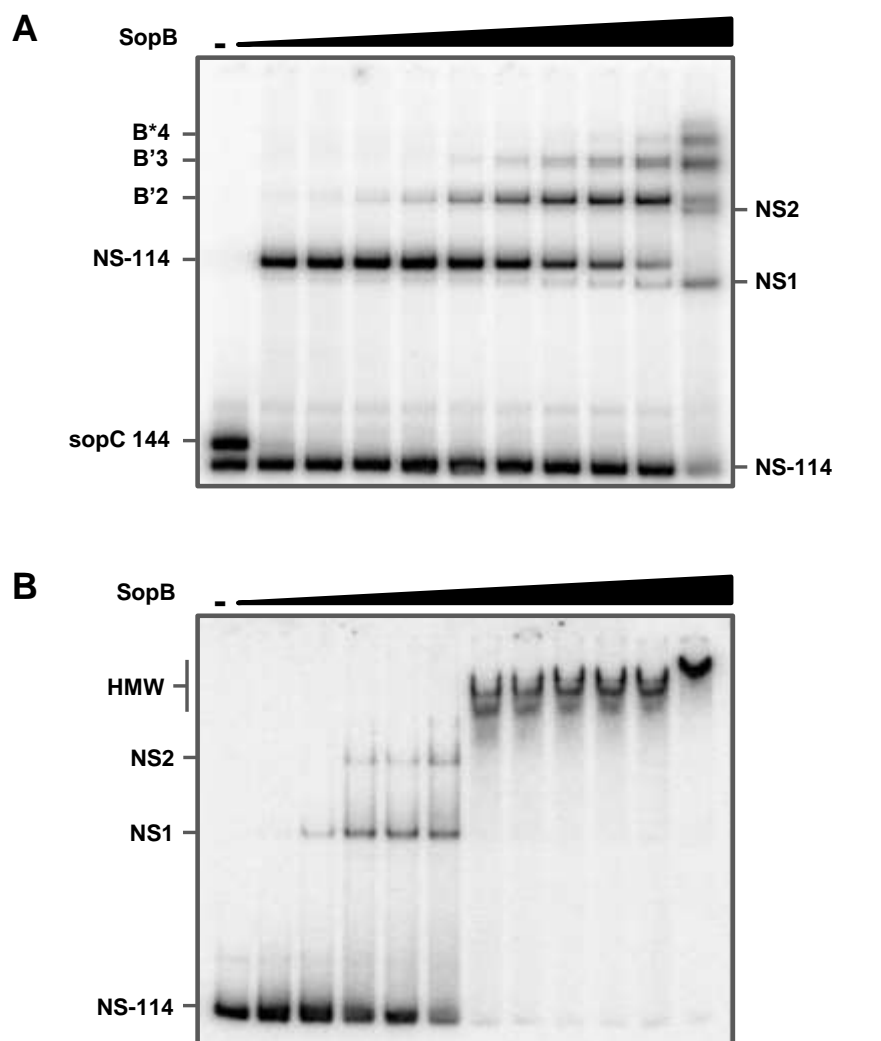
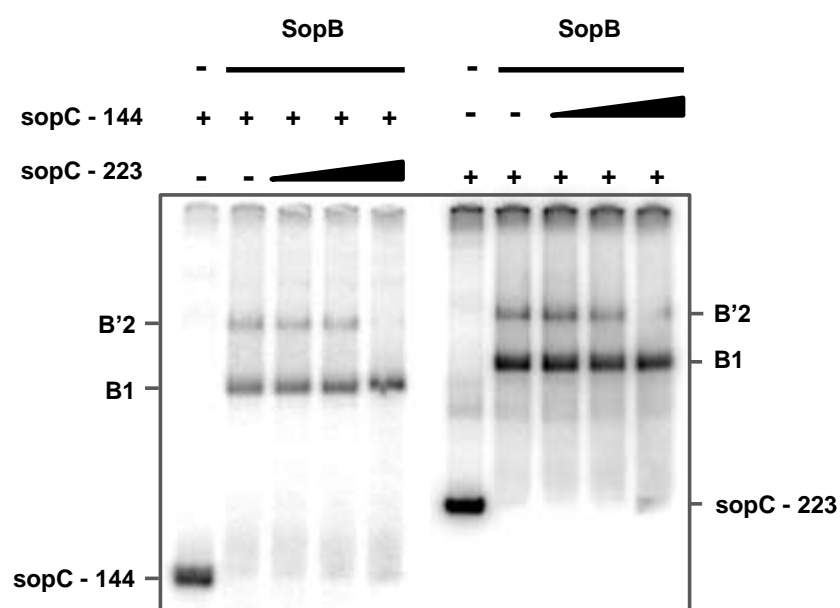


Figure 40 *Complex B'2 is not composed of two B1 complexes.*

Radiolabeled (+) *sopC*-144 (left part) or *sopC*-223 (right part) probes (0.3 nM) were incubated with SopB (450 nM) in binding buffer containing 50 mM Hepes-KOH pH 7.5, 100 mM KCl, 100 µg/ml BSA, 100 µg/ml sonicated salmon sperm DNA and 1 mM DTT. Increase concentrations (3, 30, 300 nM) of non-radiolabeled *sopC*-144 or *sopC*-223 were mixed in reactions as indicated by the black triangle. Mixtures were incubated at 30°C for 15 min and then loaded on a 6 % polyacrylamide gel in Tris-Glycine. Following electrophoresis, the gel was dried and exposed to a phosphor screen for imaging with a Typhoon (GE Healthcare). The positions of free and SopB-bound *sopC* probes are indicated on the left (*sopC*-144) and on the right (*sopC*-223). No additional band was detected upon the titration with unlabeled *sopC* fragment of different length compare to the labeled probes, indicating that B'2 is not a sandwich complex between two B1 complexes.



specific probes in the same EMSA (**Figure 39A**), showing that a non-specific DNA complex (NS1) appeared at a concentration of SopB (1 μ M) well above the 30 nM needed to detect the B'2 complex, and (ii) assaying SopB non-specific DNA binding in EMSA without an excess of non-specific competitor DNA (**Figure 39B**), showing that a complex (NS1) was observed only above 200 nM. From these results we conclude that the formation of secondary binding complexes requires other interactions than non-specific DNA binding.

We then performed EMSA with a smaller *sopC* probe (*sopC*-30) composed of 30-bp compared to the 144-bp used in the previous experiment (**Figure 38**). We found that SopB binding to *sopC*-30 (left panel) was as efficient as *sopC*-144 (right panel) for B1 complexes formation. However, no secondary B'2 complex was detected even at a high concentration of 1 μ M. This result thus indicates that flanking DNA is required for the formation of the B'2 complex. It also excludes the possibility, proposed by Schumacher et al. (2010), that each monomer of a SopB dimer could contact a different *sopC*-DNA molecule. We also performed competition experiments, with increasing concentration of unlabeled *sopC*-carrying DNA fragments of different size than the labeled ones (**Figure 40**), excluding the possibility that B'2 complexes arise from the assembly of two SopB dimers on two different large DNA molecules.

We then investigated the minimal length of the flanking DNA required for B'2 complex formation, using probes of decreasing lengths with centrally located 16-bp *sopC* sites (**Figure 38, central panels**). B'2 complexes were observed with a *sopC*-60 probe, but were barely detectable with *sopC*-40 and not detectable with *sopC*-36. The minimal length of the flanking DNA outside the 16-bp *sopC* site is thus 12-bp for B'2 complex formation (see **Table SII**). Interestingly, the longest the flanking DNA the highest the amount of B'2 complex formed, suggesting that the concentration of non-specific DNA in the close vicinity of the *sopC* binding site is important for B'2 complex formation.

All together, these results thus strongly suggest that the B'2 complex formation does not result from a unique interaction but rather requires both dimer-dimer and non-specific DNA interaction.

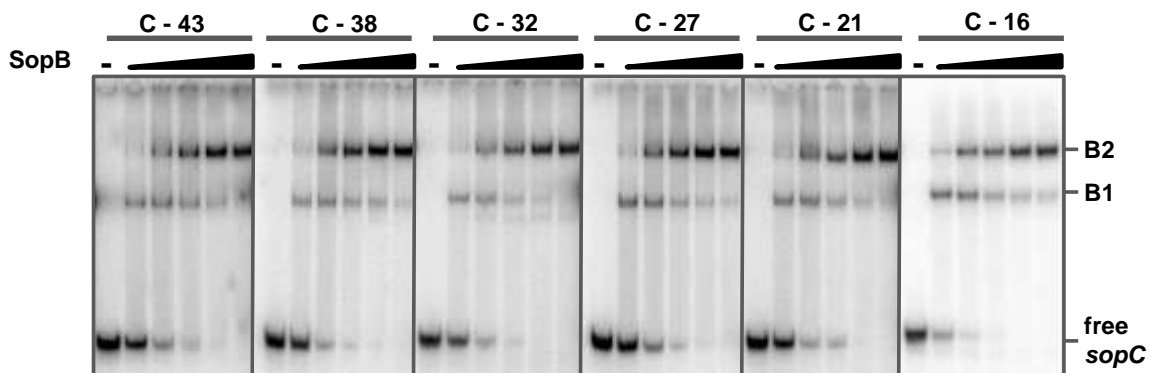
5. SopB dimers assemble every 16-bp along the DNA

SopB binding sites are regularly spaced every 43-bp in the centromere sequence (**Figure 34A**), each 16-bp site being separated by 27-bp. Note that this spacing (43-bp) corresponds to

A



B



C

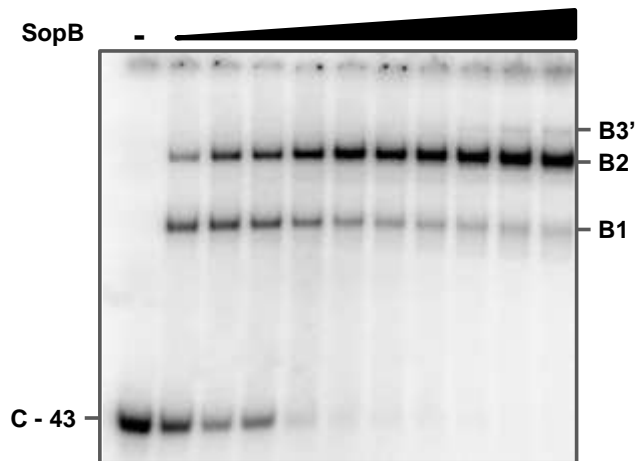


Figure 41 DNA spacing variation between two adjacent binding sites reveals that SopB dimer could bind every 16-bp without steric hindrance.

(a) Sequences and schematic representation of the *sopC* probes used in EMSA. The top sequence corresponds to the consensus sequence between two adjacent SopB binding sites. Inverted arrows indicate the 16-bp binding sites. Probes are labeled "Cn" with n corresponding to the number of nucleotide (in bp) between the first (black) and the second (grey) SopB-binding sites. On the right, the position of the SopB binding sites relative to the face of the double helix is schematically drawn. (b) EMSA with the DNA probes carrying different spacing between two adjacent SopB binding sites were performed as in **Figure 36**. SopB concentrations were 0, 10, 30, 100, 300 and 1000 nM. (c) SopB titration on C43 probe, from 10 to 5120 nM by 2-fold increments. B1 and B2 represent complexes involving one and two specific *sopC* binding sites, respectively, while B'3 corresponds to a complex involving two specific and a non-specific DNA binding sites.

four helical turns of the DNA helix (B-form) implying that SopB dimers should bind to the same face of the DNA. We wondered whether this spacing is maintained outside *sopC*. To determine the closest distance between two sites that support two SopB dimers without steric hindrance, we performed EMSA with 63-bp DNA fragments that carry two SopB binding sites; the second site being moved by steps of 5- or 6-bp relative to the first site (**Figure 41A**). With the two SopB binding sites naturally spaced by 43-bp (probe C-43), we observed two discrete complexes, B1 and B2 (**Figure 41B**, left panel); the former is more abundant at lowest SopB concentrations while the latter is more abundant at the highest as expected. This pattern was observed for all the probes tested, even for the one carrying the two sites juxtaposed without separation (probes C-16). For all probes tested, no difference in binding efficiency was observed (**Figure 41B**). These results thus clearly reveal that no steric hindrance prevent SopB to assemble every 16-bp along the DNA, a configuration on which SopB dimers would no longer be on the same face of the DNA.

The DNA fragments used in this assay should be long enough to accommodate the binding of an additional SopB dimer. We used the C-43 probe to test this possibility, and indeed we observed the detection of a third complex (B'3), corresponding to non-specific binding in between the two SopB binding sites (**Figure 41C**). In accordance to the above results (**Figure 38**), this complex is detected only at SopB concentrations above 1 μ M. At high concentration, SopB dimers can thus bind in between two specific binding sites.

6. Extended partition complexes do not assemble in stable nucleoprotein complex

To further investigate the organization of the partition complex outside of *sopC*, we performed *in vitro* footprinting assays using a 223-bp *sopC* DNA probe (**Figure 42A**). We observed a strong protection on the 16-bp inverted repeat binding site. However, no secondary footprint pattern was detected even at high SopB concentration (2 μ M) as confirmed by densitometric scanning analysis of the footprinting patterns. With a DNA substrate containing two naturally-spaced SopB binding sites (**Figure 42B**), we readily detected two protected DNA regions corresponding to both 16-bp *sopC* sites but again no protection outside or in between these specific binding sites was detected. These results are observed independently of the presence of competitor DNA (compare lanes "2000" and "2000-"). The lack of footprinting pattern corresponding to the formation of B'2 complexes is in accordance with low binding affinity outside the centromere sites.

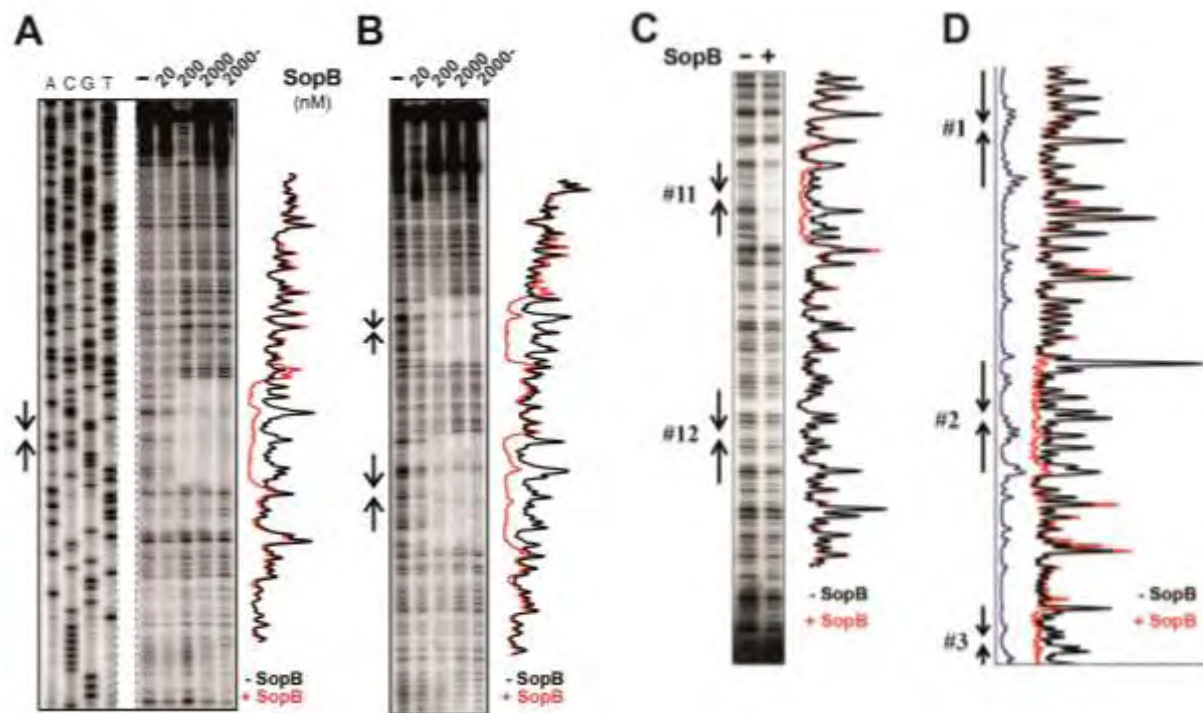


Figure 42 DNase I footprinting analyses reveal no protection pattern outside SopB-binding sites.

Footprinting assays on *sopC* DNA probes were performed *in vitro* (A, B) and *in vivo* (C, D).

(A and B) DNase I footprints on DNA fragments carrying a single (A; *sopC*-144) and two 43-bp spaced SopB-binding sites (B; *sopC*-223). SopB concentrations (nM) are indicated on top of the gels; "-" indicates that the reactions were performed in the absence of competitor DNA. Head to head arrows on the left of the gels corresponds to the 16-bp SopB binding sites. Normalized densitometric scans of the gel are shown with black and red lines corresponding to the *in vitro* footprints in the absence or in the presence of 2000 nM SopB, respectively. In (A), the left part corresponds to sequencing reactions of the *sopC*-144 probe.

(C and D) Exponentially growing DLT1472 cells carrying pZC204, induced (+) or not (-) for SopB expression, were subjected to *in situ* DNase I treatment, and total DNAs were used as templates for primer extensions. Inverted arrows correspond to the 16-bp SopB binding sites and are numbered from the SopB-proximal end as represented in **Figure 34A**. Normalized densitometric scans of the gels are shown with red and black lines corresponding to SopB induced or uninduced. In (D), only the scans are shown.

To explore the DNA protection outside of *sopC in vivo*, we performed *in situ* footprinting assays using the “drunken cell” technique (Bouet and Lane, 2009; Cassler et al. 1999). The results of the footprints and the corresponding densitometric scanning analysis (**Figure 42C**) showed that while the 11th *sopC* repeat is protected by SopB as previously observed (Bouet and Lane, 2009) the 12th repeat is not. The first *sopC* repeat is also not bound by SopB while the second is (**Figure 42D**). The first and last (12th) *sopC* repeats was previously shown to be too degenerated from the consensus sequence to be bound by SopB *in vitro* (Pillet et al. 2011). These results indicate that this is also the case *in vivo*. In addition, there is also no evidence for altered pattern in between each *sopC* site (**Figure 42C, D**). All these results thus reveal that SopB does not bind strongly enough outside *sopC* binding sites to be detected *in vitro* and *in vivo* by DNase I footprinting.

The absence of protection by SopB outside its binding site is in agreement with the whole genome ChIP-seq analysis (see above) and clearly indicates that SopB does not propagate from the centromere DNA in a polymerizing nucleofilament fashion. SopB are present at ~1000 dimers per cell in condition where F plasmids are present at ~5 copies per cell (Adachi et al. 2006; Bouet et al. 2005), i.e. ~50 SopB binding sites. SopB dimers are therefore in excess of ~20-folds. We wondered whether supernumerary SopBs are randomly localized, concentrated in the vicinity of *sopC* or even form helical structures as suggested by Adachi et al. (2006). To answer this question, we performed single molecule in single cell localization microscopy.

7. All SopB molecules are concentrated in foci at centromere sites

Previous studies using conventional epifluorescence microscopy methods showed that SopB molecules coupled to a fluorescent tag appears mostly as foci (Adachi et al. 2006; Sanchez et al. 2013), as also observed for most ParBs assembled on their cognate centromeres (Erdmann et al. 1999; Jakimowicz et al. 2005). However, the spatial resolution of these techniques is limited by the diffraction of light (~ 250 nm) and fast diffusing molecules are not detected. To further study the composition and architecture of the extended partition complexes and discriminate between complex- bound and unbound SopB we turned to single-molecule super-resolution microscopy (SMLM, Betzig et al. 2006; Hess et al. 2006). SMLM allows for the detection of single molecules and can reach higher resolutions than conventional fluorescence microscopy techniques (~30 nm). We tagged SopB with a monomeric photo-

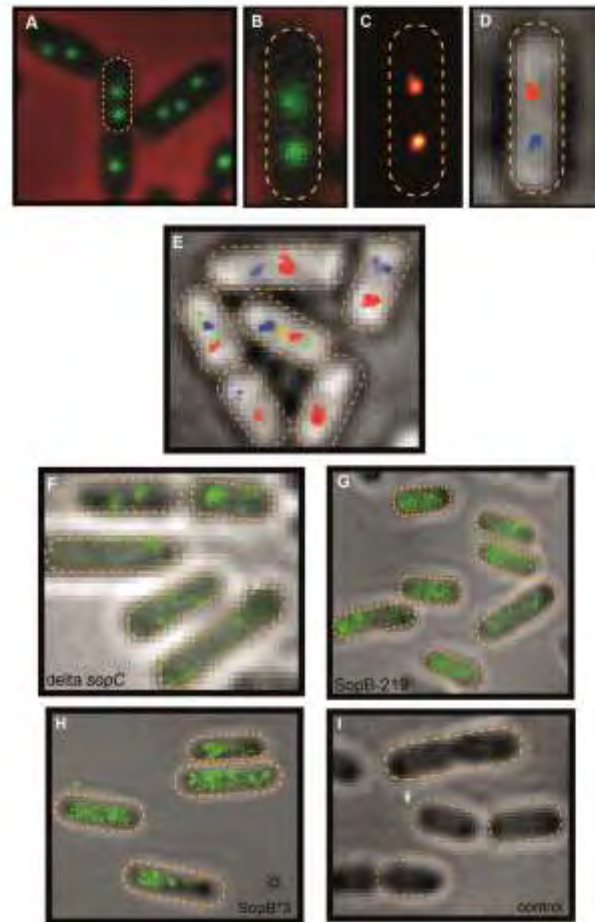


Figure 43 *SopB* assembles into compact clusters.

(A-B) Conventional epi-fluorescence image of *E. coli* cells, in which SopB (green) was tagged with (A) mVenus or (B) mEos2 and assembles in diffraction-limited foci. Membranes in both panels are shown in red.

(C) Superresolution image of an *E. coli* cell expressing SopB-mEos2 in which the orange spot represents a probability density reconstruction of the localization of SopB. (D-E) SopB clusters were automatically detected. Each single-molecule localization of the image rendered in C is represented by a single colored dot in D. Green dots indicate freely diffusing SopB (i.e. does not form part of any cluster) and blue and red dots indicate single molecule events that belong to two different clusters. A larger set of bacteria is depicted in panel E using the same color code for free SopB (green) and clusters depicted in different colors (blue, red and orange). (F-H) SopB does not form complexes in absence of *sopC*. (F) Strain of *E. coli* in which all SopB-binding sites have been removed from plasmid F shows that SopB does not form clusters. A small percentage of cells showed unspecific clusters. (G-H) SopB-mEos2 variants deficient either in (G) *sopC* binding (SopB-R219A) or in (H) DNA binding (SopB*3) do not form any clusters and SopB appears homogeneously distributed in the cytoplasm. (I) No background fluorescence is detected in bacterial cell devoid of SopB-mEos2. Scale bar (to be included), 1 μ m in all images. Orange dotted lines represent cells outlines.

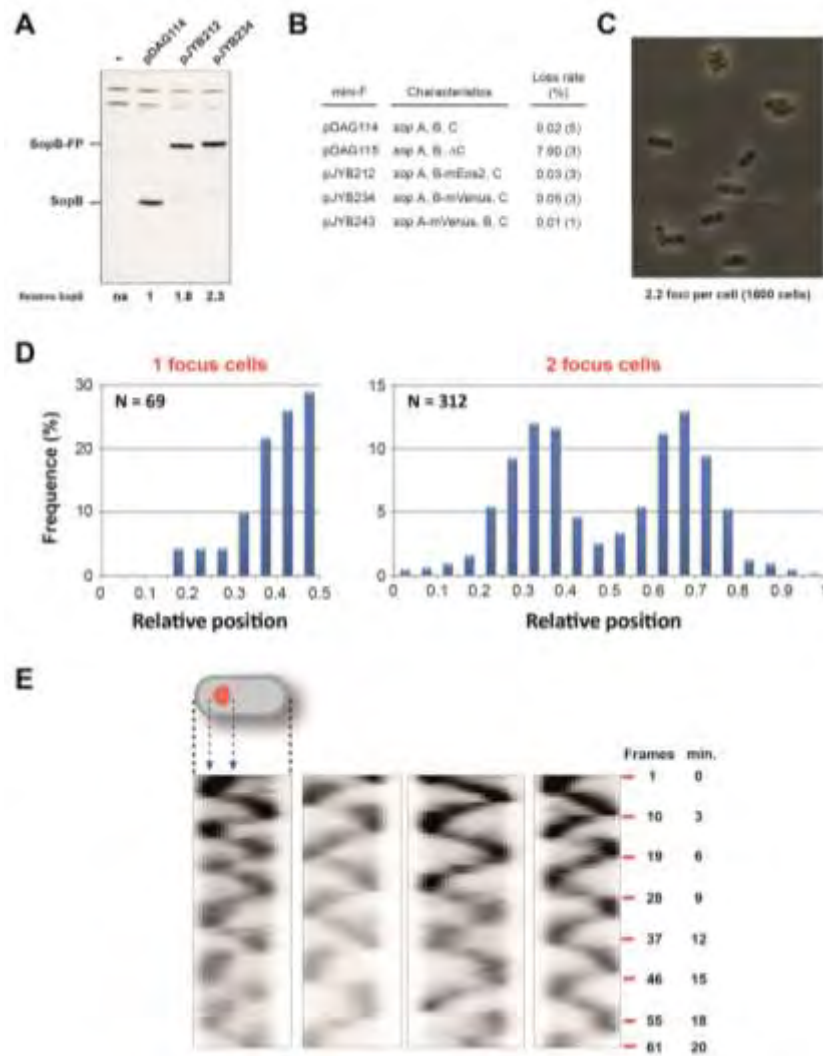


Figure 44 Expression, activity and localization of Sop proteins in condition used for epifluorescence PALM experiments.

(A) SopB-mVenus and SopB-mEos2 expression level. Samples of *E. coli* strains carrying mini-Fs with either wild type sopB (pDAG114), sopB-mEos2 (pJYB212) or sopB-mVenus (pJYB234) were harvested in exponential phase. Protein extracts were resolved on a 4-12% Bis-Tris gradient gel (Invitrogen), and the SopBs, revealed by Western blots using anti-SopB antibodies, were quantified. The level of fluorescently-tagged SopB (SopB-FP) relative to wild type SopB expressed from pDAG114 is indicated below. (B) The functionality of fluorescently-tagged SopB were assayed by plasmid stability assays as described previously (Lemonnier et al. 2000). (C) Fluorescent foci from the mini-F expressing SopB-mVenus display the expected pattern for partition complexes faithfully partitioned. Foci detection using MicrobeTracker routines indicated that cells harbor 2.2 foci per cell (average from 375 cells). (D) Intracellular positioning of SopB-mVenus foci in 1 focus (left panel) and 2 foci (right panel) containing cells as a function of the relative cell length. The number (N) of cell measured is indicated on each panel. (E) SopA-mVenus oscillations were imaged by time lapse epifluorescence microscopy. Images were taken every 20 sec for 20 min. Four representative kymographs of SopA oscillation are shown. A schematic representation of the cell with SopA-mVenus fluorescent signal (red) is presented on top of the first kymograph. The length of the each kymograph corresponds to the length of analyzed cell. For SopB-mVenus (C) and SopA-mVenus (E) imaging, cells were grown in MGC, washed in M9 and spotted on M9 agarose-coated slides, i.e. condition similar to the one used for PALM imaging.

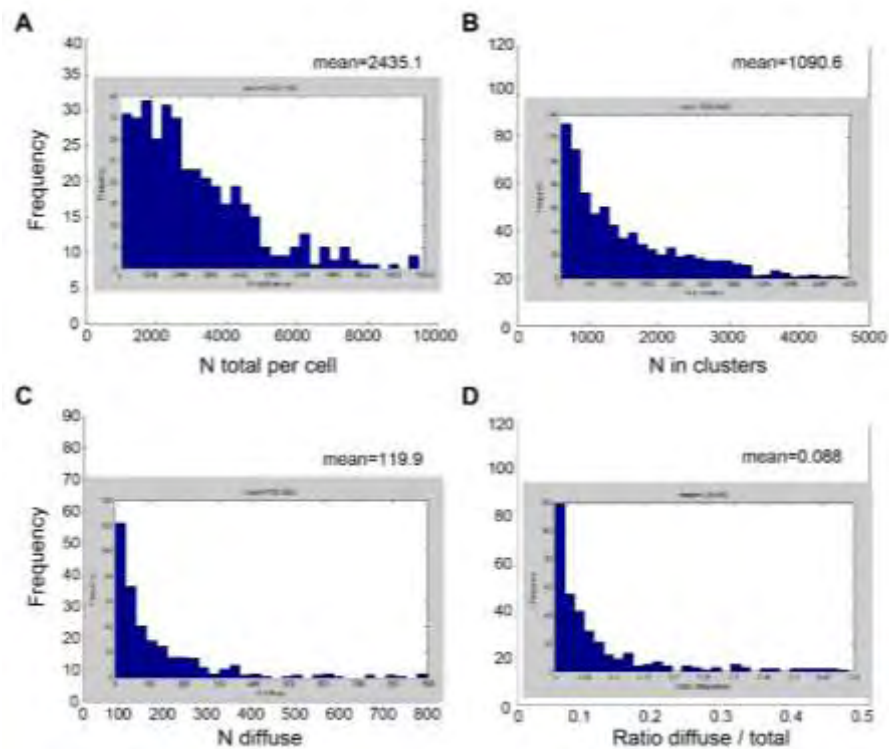


Figure 45 *The majority of SopB is recruited to the partition complexes and assembled in a small-sized nucleoprotein complex.*

(A-C) The total number of single molecule events (N) frequency distribution is plotted for (A) the whole cell, (B) individual clusters and (C) freely diffusing SopB. (D) Frequency distribution of the ratio of diffuse SopB molecules versus the total number of SopB. In average, SopB clusters contain > 90 % of the total SopB present in the cell. A Gaussian distribution was fitted to each N distribution in order to estimate the average number of events ($\langle N \rangle$). The same approach was employed to estimate the mean ratio and standard deviation of freely diffusing/cluster-located SopB.

activatable fluorescent protein (mEos2, see material and methods) and employed a microfluidics-coupled, PALM microscope to perform automatic, high-throughput super-resolution detection and analysis (Cattoni *et al.* 2012; Fiche *et al.* 2013) of SopB localization in live cells. As a control a SopB-mVenus fusion of SopB with monomeric mVenus was constructed. Both mini-Fs were fully stable and express tagged-SopBs (SopB-mEos2 or SopB-mVenus) at the same level as WT SopB (**Figure 44A**).

When imaging SopB-mVenus by epi-fluorescence, in identical experimental condition as for PALM, foci were positioned at mid-cell or at $\frac{1}{4}$, $\frac{3}{4}$ positions of the cell length (**Figure 44D**), when 1 or 2 foci were respectively observed (**Figure 44C**), with 2.2 foci per cell in average. We also verified that SopA is still oscillating two hours after cells have been loaded on the slides (**Figure 44E**). Epifluorescence visualization of SopB-mEos2 by excitation at 488 nm results in the same localization pattern (**Figure 43A**) as that observed for SopB-mVenus confirming that no mis-localization was induced by the photoactivatable tag. Next, a field of view was first imaged by bright-field microscopy and followed by PALM imaging (**see material and methods**). We observed a large number of individual SopB fluorescent traces per cell (**Figure 43D**) while practically no events were detected in the absence of mEos2 (**Figure 43I**). Interestingly, SopB mainly forms clusters (**Figure 43C-D**). To further analyze the composition of SopB foci, clusters were automatically detected from the spatial distribution of single-molecule localization events (**see material and methods**). Analysis of cluster sizes and compositions revealed that SopB assembles in compact clusters (**Figure 43C-E**) with sizes smaller than 40 nm (data not shown).

Although the number of detected single-molecule events (N) does not provide absolute protein numbers, N is still proportional to the number of single emitters, and can thus be used to obtain relative measures of the total number of proteins in cells and the proportions between freely diffusing and cluster associated proteins. When analyzing the total number of fluorescent events per cell, we quantified in average ~2400 events of SopB-mEos2 molecules (**Figure 45A**). Regarding the foci composition, we found that each cluster contains in average ~1100 fluorescent traces (**Figure 45B**), which corresponds to ~2.2 foci per cell, in complete agreement with our results in epifluorescence (see above). To discard any effect of SopB oligomerization or an undesired clustering due to the fluorescent tag, we constructed SopB-mEos2 variants that are specifically deficient either at *sopC* binding (SopB-R219A, Sanchez *et al.* 2013) or at DNA binding (SopB*3, Ah-Seng *et al.* 2009). In both cases, we observed no SopB

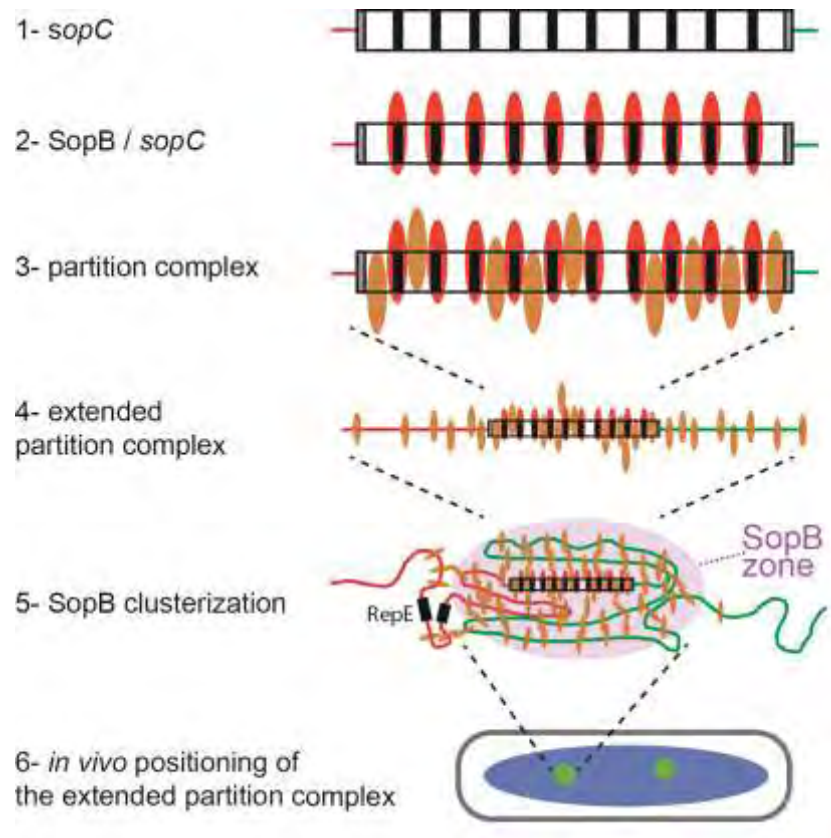


Figure 46 | Model for the extended partition complex dynamic organization.

The different steps for the ultrastructural assembly of SopB onto the F centromere and surrounding DNA are depicted from top to bottom. The F centromere *sopC* is schematically drawn (1) with the black rectangles representing the ten specific binding sites to which SopB dimers (red ovals) bind with high affinity (2). The grey rectangles correspond to the 1st and last *sopC* repeats that are shown not to be SopB-bound *in vivo* and *in vitro*. Green and red lines represent the DNA adjacent to *sopC* on the right and left sides, respectively. Through SopB-SopB and SopB-DNA interaction, SopB dimers can bind non-specifically (orange ovals) in between specifically bound-dimers to form the partition complex over the entire *sopC* sequence (3). The extended partition complex is formed by the decreasing gradient of SopB binding away from *sopC* (4). SopB-SopB contacts maintain the *sopC*-proximal DNA sequences in close vicinity and prevent SopB to diffuse away from the centromere resulting in the formation of a SopB cluster (5). The assembly of nucleoprotein complexes, such as RepE/iterons (black rectangle), that are able to pair iteron sites could take away surrounding DNA sequences and reduce their spatial proximity with the partition complex, resulting in the asymmetric distribution of SopB binding on both sides of *sopC*. The formation of highly concentrated SopB clusters enables the visualization of intense foci (green dots) when fluorescently-tagged SopB are observed in living cell (6). The blue oval and the grey flared-rectangle represent the nucleoid and the cell membrane, respectively.

clusters but an homogeneous distribution of SopB molecules all over the cytoplasm (**Figure 43G-H**). These results are inconsistent with SopB assembly in clusters being influenced by the mEos2-tag fusion (Landgraf et al. 2012), and confirm that SopB-mEos2 molecules do not self-assembled in clusters without the primary binding on the *sopC* centromere.

Finally, to estimate the proportion of freely diffusing and *sopC*-associated SopB, we quantified the number of SopB events that do not belong to foci and found that only ~120 fluorescent events per cell in average do not belong to SopB clusters (**Figure 45C**). The ratio of diffuse SopB molecules versus the total number of SopB is less than 10% (**Figure 45D**), indicating that more than 90 % of SopB molecules are clustered in a small number of foci in the cell. The intracellular localization of these SopB clusters correlates well with the localization of the F plasmid inside the cell (unpublished data).

8. Discussion

The combination of whole genome ChIP experiments, DNA binding assays and superresolution microscopy enabled us to characterize the global architecture of a type I partition complex involved in bacterial DNA segregation. The Sop partition system of the *E. coli* plasmid F belongs to the most widespread partition family and is closely related to the Par systems present on most bacterial chromosomes. It has been assumed for a long time that ParB/SopB/Spo0J centromere binding proteins spreads on DNA nearby *parS/sopC* centromere sites by side by side polymerization to form a proteo-nucleofilament outside their cognate centromere binding sites (Murray et al. 2006; Rodionov et al. 1999). However, no direct evidence for such ParB nucleoprotein complexes organization was provided for type I partition complexes. Here, we unveiled the mode by which SopB propagate along the DNA from the centromere and revealed the intracellular positioning of SopB molecules in living cells at the single molecule level. Our results showed that most SopB molecules are highly clustered at the centromere site and that SopB does not spread by polymerization along the DNA but rather propagates by stochastic binding in the close vicinity of the centromere binding site (**see model Figure 46**). These findings have major implications for the active segregation mechanism in bacteria.

Using ChIP analysis we have shown that SopB proteins propagate on both sides of the *sopC* centromere site over ~20 Kb (**Figure 36**). However, our results exclude the possibility that SopB spreads by polymerization along the DNA outside the centromere for three main reasons.

First, the rapid decrease in the number of SopB bound-sequences over only 20-bp on both sides of *sopC* (**Figure 36C**) indicates that the SopB ultrastructural architecture just outside the centromere is labile. It is therefore not compatible with the formation of SopB polymers along the DNA for which we would expect a much slower decrease in the close vicinity of the centromere. Second, in the case of SopB polymers stabilized by side by side interactions, we would have expected to detect DNA-footprints nearby the *sopC* site. This is not the case as no sign of protection on centromere-flanking DNA, and even in between *sopC* sites, was observed both *in vitro* and *in vivo* (**Figure 42**). Third, the intracellular level of SopB appears too low to support formation of long SopB polymers. SopB has been estimated to be present at ~1000 dimers per cell (Adachi et al. 2006; Bouet et al. 2005), condition where the mini-F is present at ~5 copies in average (Bouet et al. 2005), i.e. ~50 SopB specific binding sites per cell. We demonstrated that SopB dimers could bind every 16-bp without steric hindrance (**Figure 41B**). Assuming that SopB are assembled every 16-bp in a SopB-nucleofilament manner then SopB dimers would cover 16 Kb. This corresponds to ~3 Kb per plasmid if distributed homogenously on each copy present in a cell, i.e. ~1.5 Kb of spreading capacity on both sides of *sopC*. This is well below the ~20 Kb detected *in vivo* and thus implies that SopB could not extent as long polymers along the DNA adjacent to centromeres. Together, all these findings thus indicate that SopB does not bind outside *sopC* with a regular pattern.

The basis for type I partition complex architecture based on ParB polymerization along the DNA arises from experiments suggesting the existence of road blocks that prevent ParB binding to extend beyond them. For instance, ParB P1 was not detected on DNA sequence after the iteron loci on which the RepA initiator proteins bind tightly (Rodionov et al. 1999). As for plasmid P1, the plasmid F carries two iteron regions (*ori2* iterons and *incC*, Uga et al. 1999). These regions can interact together through the RepE protein and regulate the initiation of DNA replication by a handcuffing mechanism (Das and Chattoraj, 2004). Our highly sensitive ChIP results show that SopB F propagation is not blocked by the iterons (**Figure 36C**) indicating that the RepE/iterons nucleoprotein complexes are not acting as physical barriers. However, a careful analysis of the SopB-bound on the left side of the centromere indicates an important change in the distribution of SopB in the close vicinity of the *incC* site, as revealed by a different slope in the number of reads (**Figure 36**). This change in the "regime" by which SopB propagates on the DNA indicates that the RepE/*incC* assembly has an influence on SopB propagation but does not prevent it as a barrier would do. One possibility could be that the RepE/*incC* assembly

reduces the spatial proximity between *sopC* and sequences beyond the iterons, thus reducing the probability that SopB dimers interact with these sequences. Any case, other analyses will be required to precisely define the local effect of a nucleoprotein assembly on the propagation of SopB.

Collectively, our data are in favor of a new model involving the stochastic binding of SopB on non-specific DNA sequence in the close vicinity of the primary SopB assembly sites. In agreement with this hypothesis, we found that the longer the DNA is in the vicinity of *sopC* sites, the more secondary complexes are formed (**Figure 38**). **Figure 46** recapitulates the different steps involved in the assembly of the extended partition complexes. We propose that the strong and specific binding of SopB exclusively on *sopC* sites (Pillet et al. 2011) nucleates the formation of highly concentrated clusters of SopB molecules by the synergistic combination of the relatively weak SopB-SopB and SopB-DNA interactions. In our model, the DNA adjacent to *sopC* provides the lattice for caging SopB dimers and thus preventing their escape, as observed using *in vivo* single SopB molecule imaging (**Figure 43 and Figure 45**).

Active segregation by type **II** and type **III** partition systems is also mediated by the assembly of centromere binding proteins (ParR) on their cognate centromere sites (*parC*). The architecture of these ParRC assemblies is well-known (Aylett and Lowe, 2012; Salje and Lowe, 2008), with the centromere DNA wrapped around a helical scaffold of ParBs. By contrast to these types for which the centromere organization is very similar - direct DNA repeats of regularly spaced ParR-binding sites - the organization of type **I** centromeres is highly different for one system to another (Hayes and Barilla, 2006). Type **I** centromeres consist of either direct repeats, inverted repeats or both with or without internal symmetry. The repeats can be extensive in a single locus or distributed over relatively large region as in the case of chromosomal *parS* sites. For instance, the plasmid P1 centromere site is a complex arrangement of two binding motifs present several times over 85-bp and separated by a strong IHF bending site (Funnell, 1991; Schumacher and Funnell, 2005). Centromere sites present on bacterial chromosomes are much simpler, composed of several 16-bp *parS* sites present usually in a few copies (2-4) over several Kb (Livny et al. 2007). The common features of all these type **I** centromere sites are the strong primary binding of ParBs on *parS* and the subsequent propagation of ParB along the adjacent DNA. The precise role of this architectural organization is currently unknown.

For the first time, we observed the localization of all individual SopB molecules in the single cell using super-resolution microscopy. It has been previously described that some SopB co-localized with SopA in a helix pattern (Adachi *et al.* 2006). Our results are not consistent with this observation as most of the SopB molecules are highly concentrated in only a small number of foci. Less than 10% of SopB are not belonging to clusters. These clusters correspond to the number of SopB foci observed in epifluorescence microscopy and are positioned at the same place as the plasmids F (**Figure 44C, D**). Our results clearly indicate that most SopB molecules are not moving along with the SopA oscillating structure in the cytoplasm (**Figure 43E**). This finding is of particular importance for the modeling of the type I partition mechanism, and models with the assumption that ParB/SopB are traveling with ParA/SopA should be revisited. This result is also in agreement with our recent finding showing that the stimulation of SopA ATP hydrolysis, which requires SopB and DNA, is not involved in the oscillation behavior of SopA (Ah-Seng *et al.* 2013), i.e. when SopA is located at the nucleoid hedge the presence of SopB in its close vicinity is not required to initiate SopA relocation on the other side even though both SopB and *sopC* are needed for SopA dynamics. The concentration of most SopB molecules at centromere sites appears important for SopA dynamics. We propose that SopB concentration around the centromere DNA forms a higher order structure that is crucial to its correct function in the partition process. Although our results shed light in the partition process, more work is still requested to decipher at the molecular level the type I-mediated DNA segregation mechanism.

D. Material and methods

1. Bacterial strains, plasmids and growth conditions

Strains are derivatives of *Escherichia coli* K-12 and are listed, together with plasmids, in Table SIII. Constructions of the mini-F plasmids carrying fluorescent tag fused to *sopA* or *sopB* genes were made using the In-Fusion kit (Clontech). The plasmid F is a pOX38 derivative (Guyer *et al.* 1981) in which the *ccdB* toxin was replaced with the *cat* gene by lambda-red recombineering (gift from A. wright). The mVenus-Yfp gene was constructed by introducing the monomeric A207K mutation (Zacharias *et al.* 2002) in the venus-Yfp gene by Quik Change mutagenesis. All constructs were verified by DNA sequencing (MWG). Cultures were essentially grown at 37°C with aeration in LB (Miller, 1972) containing thymine (10 µg.ml⁻¹) and chloramphenicol (10 µg.ml⁻¹) as appropriate. LB was supplemented with 1.5% agar (Difco) for

202

solid medium. For microscopy assay, cultures were grown at 30°C with aeration in MGC (M9 minimal medium supplemented with glucose (0.4 %), casamino acids (0.2 %), MgSO₄ (1 mM), CaCl₂ (0.1 mM), thiamine (1 µg.ml⁻¹), leucine (20 µg.ml⁻¹) and thymine (40 µg.ml⁻¹)).

2. Chromatin immuno-precipitation

The ChIP procedure was adapted from (Cho *et al.* 2011) with the following modifications. Strains DLT1215 carrying pDAG114 or pOX38B were grown overnight in LB, diluted 1/200 into 50 ml fresh LB and grown at 37°C to OD₆₀₀ ~0.6. Cultures (40 ml) were treated with formaldehyde (1% final concentration), incubated at room temperature for 30 min with gentle agitation, and quenched by addition of 0.5 M glycine for 15 min with agitation. Cells were harvested by centrifugation (6 krpm for 10 min. at 4 °C), washed twice in Tris-buffered saline (20mM Tris pH7.6, 100 mM NaCl) and resuspended in 0.5 ml lysis buffer (10mM Tris, pH 7.8, 10mM EDTA, 100 mM NaCl, 20% sucrose). Lysozyme (1 mg.ml⁻¹) was added and reactions were incubated for 30 min at 30°C. Subsequently, 0.5 ml of cold 2x IP buffer (50mM Hepes-KOH, pH 7.5, 150mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS) was added and reactions were chilled on ice. The lysates were then sonicated 18 times for 30 sec (maximum power) with pulsed 50% duty cycle using Sonicator Bioruptor® Plus (Diagenode). Cell debris was removed by centrifugation at 20,000 x g for 30 min at 4°C. Aliquots of each sample were frozen at -20°C. Protein A-coupled Magnetic beads (Bio-Adembeads, Ademtech) were incubated with a blocking buffer (0.1 mg.ml⁻¹ BSA, 1 mg.ml⁻¹ tRNA) for 4 h at 4°C. Protein A beads (25 µl) were added to 500 µl of cell lysates, incubated for 1 h at 4°C with gentle agitation 25 µl, and removed using the Adem-Mag Modulo (Ademtech) after the incubation. Affinity-purified anti-SopB polyclonal rabbit antibody (7 µg) was added to each "pre-cleared" sample and incubated overnight at 4°C with gentle agitation. Blocked protein A beads (25 µl) were added to each sample and incubated for 2 hours at 4°C with gentle agitation. Beads were collected and washed once with 2X IP buffer, twice with buffer WI (50mM Hepes-KOH, pH 7.5, 500mM NaCl, 1mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS), once with buffer WII (10 mM Tris-HCl pH 7.8, 250 mM LiCl, 1 mM EDTA, 0.5% Nonidet-P40, 0.1% sodium deoxycholate), once with TE buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA). DNA and proteins were eluted from the beads by resuspension in 300 µl of elution buffer (50mM Tris-HCl at pH 7.5, 10mM EDTA, 1% SDS) complemented with RNase A (10 µg) and incubation for 1 h at 37°C. Proteinase K (0.3 mg.ml⁻¹) was then added and samples were incubated for 2 h at 37°C. Samples

for "Input control" were thawed, treated as above with RNase A for 1 h at 37°C then with proteinase K (0.3 mg.ml⁻¹) and 60 µl of 10% SDS for 2 h at 37°C. Formaldehyde cross-linking was reversed by overnight incubation at 65°C. DNA in input control and ChIP samples was purified using phenol-chloroform extraction and resuspended after EtOH precipitation in 100 µl H₂O.

3. q-ChIP analysis

Precipitated DNA from ChIP preparation were analyzed by real-time PCR in parallel with input genomic DNA as a standard, with primer pairs (listed in Table S1), using a Roche Light Cycler and a Light Cycler FastStart DNA Master SYBR green kit. ChIP enrichment values represent the relative abundance of a sequence of interest compared to a control of immunoprecipitation without Anti-SopB antibody. All enrichment values represent the average of at least three independent biological replicates.

4. ChIP-seq

This work has benefited from the facilities and expertise of the high throughput sequencing platform of IMAGIF (Centre de Recherche de Gif - www.imagif.cnrs.fr)

5. Electromobility shift assay

Experiments were performed as described (Bouet et al. 2007), using 0.3 or 1 nM radiolabeled DNA probes in the presence of sonicated salmon sperm DNA as competitor (100 µg.ml⁻¹). Binding reaction products were resolved at 4°C by electrophoresis on 6% polyacrylamide gels in TGE (25 mM Tris base, 190 mM glycine, 1 mM EDTA) for 2-4 h at 180 V. DNA probes were generated by PCR DNA fragments or by annealing PAGE-purified complementary oligonucleotides (Sigma-Aldrich).

6. Footprint assays

In vitro DNase I Footprinting experiments were performed using 10 nM ³²P- labelled DNA carrying a single 16-bp SopB binding site (*sopC*-144) or two binding site (*sopC*-223). Reaction mixtures were assembled on ice in 20 µl volume containing 30 mM HEPES-KOH (pH 7.5), 150 mM KCl, 10% glycerol, 100 µg.ml⁻¹ BSA, 100 µg.ml⁻¹ sonicated salmon sperm DNA and 1 mM DTT, and were incubated for 15 min. at 30°C. 2 µl of 1 U.ml⁻¹ DNase I solution (BioLabs) was then added and reactions were incubated for 1 min at 37°C. Reactions were stopped by adding 5µl cold 0.5 M EDTA and EtOH precipitated. DNA was resuspended in 10 µl H₂O.

In situ DNase I Footprinting was performed as previously described (Bouet and Lane, 2009). Briefly, strain DLT1472 (lacOP::sopB; Bouet et al. 2005) transformed with pZC204, a p15A derivative carrying the full length *sopC*, was grown overnight in LB containing chloramphenicol, diluted 2000-fold into the same medium with or without IPTG (250 μ M) and incubated at 37 °C. At OD₆₀₀ ~0.2, 3 ml samples were permeabilized with cold ethanol and subjected to DNase I footprinting prior to extraction and purification of the DNA (final volume 50 μ l of TE). Sixteen cycles of primer extension were then carried out using 5.5 μ l of DNA in a 10 μ l volume containing thermostable Taq DNA polymerase with 1x Thermo Pol buffer (BioLabs) and radiolabeled oligonucleotides DCF1 (5'-ATGTCCTTTGTTACAGGCCAGAAAGC-3') or DCF4 (5'-AGGCATACAACCGCACACCGAGGTTAC-3'). 5 μ l of 95% formamide, 10 mM Na₃EDTA, 0.02% bromophenol blue, 10 mM NaOH were added to primer extension reactions.

In vitro and *in situ* DNaseI reaction mixtures (3-5 μ l) were incubated at 95 °C for 5 min prior fractionation by electrophoresis on a 6% polyacrylamide-urea gel. The DNase I patterns were revealed by phosphorimaging and scanned using ImageGauge software (Fuji).

7. PhotoActivated Localization Microscopy and epifluorescence microscopy

Cells were cultured overnight at 30 °C with aeration in MGC media. An exponentially growing dilution (OD ~ 0.3-0.5) was then diluted to OD ~ 0.05 in MGC and incubated at 30°C. When reaching OD ~0.4, bacteria were centrifuged, pelleted and resuspended in M9 without glucose and CSA (hereafter M9) and containing 40 nm fluorescent beads (Invitrogen, France) used as fiducial marks. Bacterial resuspension was then injected into microfluidic chamber preincubated with 0.01% (v/v) poly-L-Lysine and incubated for 5 minutes to allow adherence of bacteria to the poly-L-lysine coated surface. A high flow force (~ 200 μ L/s of M9 media) was applied to flatten cells against the surface. Images were acquired at 19°C to slow cell movement and morphological changes.

PALM imaging and analysis was performed as described (Cattoni et al. 2013; Fiche et al. 2013). Briefly, cells injected into the microfluidic chambers were first imaged by bright-field microscopy and followed by PALM imaging. PALM was performed by continuously acquiring ~30000 images at 20Hz (55 ms per frame) under continuous illumination with a 561 nm read-out laser and by applying 405nm pulses for photo-activation. The length and intensity of UV pulses were modified during the course of the experiment to maintain the density of activated fluorophores constant while ensuring that only one protein is activated at a time in a single

diffraction-limited spot. After imaging was completed, bacteria were automatically detected on the bright field image using a modified version of MicrobeTracker (Sliusarenko et al. 2011). Single-molecule localizations were obtained by using MTT (Serge et al. 2008). Bacterial contours and localization coordinates were further processed using PALMcbcs (Fiche et al. 2013). Emission from fluorescent beads was used to correct for sample drift and chromatic aberrations between bright field and fluorescent protein images. Additionally, PALMcbcs was used to automatically classify cells according to its size and analyze the distribution of localization events to automatically detect and classify SopB clusters. Then, for each individual cluster, we calculated its size, number of events and cellular localization.

Acknowledgements

We thank S. Depaul for technical assistance with stability assays and media preparation, I. Canal for sequencing gels preparation, N. Campo for helpful advices on epifluorescence microscopy, A. Wright for providing the pOX38-B version of plasmid F, and all members of the “segregation” team for discussions. We are grateful to F. Cornet for critical reading of the manuscript.

Funding

This work was supported by the Agence National pour la Recherche (2010 BLAN 1316 01). AS was supported by a bourse de thèse du Ministère de la Recherche et de l’Enseignement, and by the Fondation pour la Recherche Médical (FDT20130928025). Fundings for microscope equipments were provided by the Fonds Européen de Développement Régional and by the Association pour le Recherche sur le Cancer.

E. Bibliography

- Adachi, S., Hori, K., and Hiraga, S. (2006). Subcellular positioning of F plasmid mediated by dynamic localization of SopA and SopB. *J Mol Biol* 356, 850-863.
- Ah-Seng, Y., Lopez, F., Pasta, F., Lane, D., and Bouet, J.Y. (2009). Dual role of DNA in regulating ATP hydrolysis by the SopA partition protein. *J Biol Chem* 284, 30067-30075.
- Ah-Seng, Y., Rech, J., Lane, D., and Bouet, J.Y. (2013). Defining the role of ATP hydrolysis in mitotic segregation of bacterial plasmids. *PLoS genetics* 9.

- Aylett, C.H., and Lowe, J. (2012). Superstructure of the centromeric complex of TubZRC plasmid partitioning systems. *Proc Natl Acad Sci U S A*.
- Betzig, E., Patterson, G.H., Sougrat, R., Lindwasser, O.W., Olenych, S., Bonifacio, J.S., Davidson, M.W., Lippincott-Schwartz, J., and Hess, H.F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science* 313, 1642-1645.
- Biek, D.P., and Shi, J.P. (1994). A single 43-bp *sopC* repeat of plasmid mini-F is sufficient to allow assembly of a functional nucleoprotein partition complex. *Proc Natl Acad Sci USA* 91, 8027-8031.
- Bouet, J.Y., Ah-Seng, Y., Benmeradi, N., and Lane, D. (2007). Polymerization of SopA partition ATPase: regulation by DNA binding and SopB. *Mol Microbiol* 63, 468-481.
- Bouet, J.Y., Campo, N.J., Krisch, H.M., and Louarn, J.M. (1996). The effects on *Escherichia coli* of expression of the cloned bacteriophage T4 nucleoid disruption (*ndd*) gene. *Mol Microbiol* 20, 519-528.
- Bouet, J.Y., and Funnell, B.E. (1999). P1 ParA interacts with the P1 partition complex at *parS* and an ATP-ADP switch controls ParA activities. *EMBO J* 18, 1415-1424.
- Bouet, J.Y., and Lane, D. (2009). Molecular basis of the supercoil deficit induced by the mini-F plasmid partition complex. *J Biol Chem* 284, 165-173.
- Bouet, J.Y., Rech, J., Egloff, S., Biek, D.P., and Lane, D. (2005). Probing plasmid partition with centromere-based incompatibility. *Mol Microbiol* 55, 511-525.
- Bouet, J.Y., Surtees, J.A., and Funnell, B.E. (2000). Stoichiometry of P1 plasmid partition complexes. *J Biol Chem* 275, 8213-8219.
- Breier, A.M., and Grossman, A.D. (2007). Whole-genome analysis of the chromosome partitioning and sporulation protein Spo0J (ParB) reveals spreading and origin-distal sites on the *Bacillus subtilis* chromosome. *Mol Microbiol* 64, 703-718.
- Cassler, M.R., Grimwade, J.E., McGarry, K.C., Mott, R.T., and Leonard, A.C. (1999). Drunken-cell footprints: nuclease treatment of ethanol-permeabilized bacteria reveals an initiation-like nucleoprotein complex in stationary phase replication origins. *Nucleic Acids Res* 27, 4570-4576.
- Cattoni, D., Fiche, J., and Nollmann, M. (2012). Single-molecule super-resolution imaging in bacteria. *Curr Opin Microbiol* 15, 758-763.
- Cattoni, D.I., Chara, O., Godefroy, C., Margeat, E., Trigueros, S., Milhiet, P.E., and Nollmann, M. (2013). SpoIII^E mechanism of directional translocation involves target search coupled to sequence-dependent motor stimulation. *EMBO Rep* 14, 473-479.

-
- Cho, H., McManus, H.R., Dove, S.L., and Bernhardt, T.G. (2011). Nucleoid occlusion factor SlmA is a DNA-activated FtsZ polymerization antagonist. *Proc Natl Acad Sci U S A*.
- Das, N., and Chattoraj, D.K. (2004). Origin pairing ('handcuffing') and unpairing in the control of P1 plasmid replication. *Molecular microbiology* 54, 836-849.
- Erdmann, N., Petroff, T., and Funnell, B.E. (1999). Intracellular localization of P1 ParB protein depends on ParA and *parS*. *Proc Natl Acad Sci USA* 96, 14905-14910.
- Fiche, J.B., Cattoni, D.I., Diekmann, N., Langerak, J.M., Clerte, C., Royer, C.A., Margeat, E., Doan, T., and Nollmann, M. (2013). Recruitment, Assembly, and Molecular Architecture of the SpoIIIE DNA Pump Revealed by Superresolution Microscopy. *PLoS Biol* 11, e1001557.
- Funnell, B.E. (1991). The P1 plasmid partition complex at *parS*: the influence of *Escherichia coli* integration host factor and of substrate topology. *J Biol Chem* 266, 14328-14337.
- Gerdes, K., Howard, M., and Szardenings, F. (2010). Pushing and pulling in prokaryotic DNA segregation. *Cell* 141, 927-942.
- Guyer, M.S., Reed, R.R., Steitz, J.A., and Low, K.B. (1981). Identification of a sex-factor-affinity site in *E. coli* as gamma delta. *Cold Spring Harbor symposia on quantitative biology* 45 Pt 1, 135-140.
- Hanai, R., Liu, R.P., Benedetti, P., Caron, P.R., Lynch, A.S., and Wang, J.C. (1996). Molecular dissection of a protein SopB essential for *Escherichia coli* F plasmid partition. *J BiolChem* 271, 17469-17475.
- Hayes, F., and Barilla, D. (2006). The bacterial segrosome: a dynamic nucleoprotein machine for DNA trafficking and segregation. *Nat Rev Microbiol* 4, 133-143.
- Hess, S.T., Girirajan, T.P., and Mason, M.D. (2006). Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophys J* 91, 4258-4272.
- Jakimowicz, D., Gust, B., Zakrzewska-Czerwinska, J., and Chater, K.F. (2005). Developmental-stage-specific assembly of ParB complexes in *Streptomyces coelicolor* hyphae. *J Bacteriol* 187, 3572-3580.
- Kusiak, M., Gapczynska, A., Plochcka, D., Thomas, C.M., and Jagura-Burdzy, G. (2011). ParB binding and spreading on DNA determine its biological function in *Pseudomonas aeruginosa*. *J Bacteriol*.
- Landgraf, D., Okumus, B., Chien, P., Baker, T.A., and Paulsson, J. (2012). Segregation of molecules at cell division reveals native protein localization. *Nat Methods* 9, 480-482.

- Lemonnier, M., Bouet, J.Y., Libante, V., and Lane, D. (2000). Disruption of the F plasmid partition complex *in vivo* by partition protein SopA. *Mol Microbiol* 38, 493-505.
- Lin, D.C.H., and Grossman, A.D. (1998). Identification and characterization of a bacterial chromosome partitioning site. *Cell* 92, 675-685.
- Livny, J., Yamaichi, Y., and Waldor, M.K. (2007). Distribution of centromere-like *parS* sites in bacteria: insights from comparative genomics. *J Bacteriol* 189, 8693-8703.
- Lynch, A.S., and Wang, J.C. (1995). SopB protein-mediated silencing of genes linked to the *sopC* locus of *Escherichia coli* F plasmid. *Proc Natl Acad Sci USA* 92, 1896-1900.
- Moller-Jensen, J., Ringgaard, S., Mercogliano, C.P., Gerdes, K., and Lowe, J. (2007). Structural analysis of the ParR/*parC* plasmid partition complex. *Embo J*.
- Murray, H., Ferreira, H., and Errington, J. (2006). The bacterial chromosome segregation protein Spo0J spreads along DNA from *parS* nucleation sites. *Mol Microbiol* 61, 1352-1361.
- Pillet, F., Sanchez, A., Lane, D., Anton Leberre, V., and Bouet, J.Y. (2011). Centromere binding specificity in assembly of the F plasmid partition complex. *Nucleic Acids Res* 39, 7477-7486.
- Pratto, F., Suzuki, Y., Takeyasu, K., and Alonso, J.C. (2009). Single-molecule analysis of protein DNA complexes formed during partition of newly replicated plasmid molecules in *Streptococcus pyogenes*. *J Biol Chem* 284, 30298-30306.
- Rodionov, O., Lobočka, M., and Yarmolinsky, M. (1999). Silencing of genes flanking the P1 plasmid centromere. *Science* 283, 546-549.
- Salje, J., and Lowe, J. (2008). Bacterial actin: architecture of the ParMRC plasmid DNA partitioning complex. *Embo J* 27, 2230-2238.
- Sanchez, A., Rech, J., Gasc, C., and Bouet, J.Y. (2013). Insight into centromere-binding properties of ParB proteins: a secondary binding motif is essential for bacterial genome maintenance. *Nucleic Acids Res* 41, 3094-3103.
- Schumacher, M.A., and Funnell, B.E. (2005). Structures of ParB bound to DNA reveal mechanism of partition complex formation. *Nature* 438, 516-519.
- Schumacher, M.A., Glover, T.C., Brzoska, A.J., Jensen, S.O., Dunham, T.D., Skurray, R.A., and Firth, N. (2007a). Segrosome structure revealed by a complex of ParR with centromere DNA. *Nature* 450, 1268-1271.
- Schumacher, M.A., Mansoor, A., and Funnell, B.E. (2007b). Structure of a four-way bridged ParB-DNA complex provides insight into P1 segrosome assembly. *J Biol Chem*.

Serge, A., Bertaux, N., Rigneault, H., and Marguet, D. (2008). Dynamic multiple-target tracing to probe spatiotemporal cartography of cell membranes. *Nat Methods* 5, 687-694.

Sliusarenko, O., Heinritz, J., Emonet, T., and Jacobs-Wagner, C. (2011). High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. *Mol Microbiol* 80, 612-627.

Uga, H., Matsunaga, F., and Wada, C. (1999). Regulation of DNA replication by iterons: an interaction between the *ori2* and *incC* regions mediated by RepE-bound iterons inhibits DNA replication of mini-F plasmid in *Escherichia coli*. *Embo J* 18, 3856-3867.

Vecchiarelli, A.G., Mizuuchi, K., and Funnell, B.E. (2012). Surfing biological surfaces: exploiting the nucleoid for partition and transport in bacteria. *Mol Microbiol* 86, 513-523.

Wu, M., Zampini, M., Bussiek, M., Hoischen, C., Diekmann, S., and Hayes, F. (2011). Segrosome assembly at the pliable *parH* centromere. *Nucleic acids research*.

Yates, P., Lane, D., and Biek, D.P. (1999). The F plasmid centromere, *sopC*, is required for full repression of the *sopAB* operon. *J Mol Biol* 290, 627-638.

Zacharias, D.A., Violin, J.D., Newton, A.C., and Tsien, R.Y. (2002). Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* 296, 913-916.

F. Supplementary data

Table SI: Primer pairs used for quantitative PCR assays.

| Names | Sequences right (R) | Sequences left (L) | used with |
|-------|-----------------------|-----------------------|------------------|
| AS18 | GAACAGTGGGCCAGAGAGA | GAAAAGCCAGCACCTGATG | pOX38-B, pDAG114 |
| AS19 | CTCCTACGCGCGCAATTAAC | GGGAACCTGTCTGCAATGA | pOX38-B, pDAG114 |
| AS20 | ACCCGTCCTCGAATTTCCAT | GGCGCTGGAGAATAGGTGAA | pOX38-B, pDAG114 |
| AS21 | CATGGCATACGGATTGGTGA | CGGCGATGAAAAAGGCTATG | pOX38-B, pDAG114 |
| AS22 | ACAGCAGCAATCCACAGCAG | TCCATCAATGCCTGTCAAGG | pOX38-B, pDAG114 |
| AS23 | GAGGACGCCGAATTTTAGGG | TAAGCGGGCAAAGTCGGTTA | pDAG114 |
| AS24 | CAGGTTTCATCATGCCGTTTG | AGGCGTTTAAGGGCACCAAT | pDAG114 |
| AS26 | ACGGCGGCTTAATACTGCAA | GGCCTTGTTGGGTATGAGGTG | pOX38-B |
| AS27 | ATCGGAATCAAAGCGACAGC | CGGAGGCTGCGTAAATGAAC | pOX38-B |
| AS28 | CGTGGTACAGCCACAACCTGG | GCGACCGTTTTTCATCTGGTC | pOX38-B |
| AS29 | GTGCAGTTCTTCTGCGGTCA | ACTGTTCCCTCAGCACAG | pOX38-B |
| AS30 | AAGCCCATCGTGGTGTCAAT | GCACTGAAACTGCCTGACGA | pOX38-B |
| AS 35 | AGCATGTGCATCGCATAGGA | AACGGTCAGGACCTGGATTG | pDAG114 |
| AS 36 | ACACGAAGCCCCAGTTCAAG | TTGAGAAGCGTCGACAGGAA | pOX38-B |

| | | | |
|-------|----------------------|----------------------|---------------------------|
| AS 38 | TGGGAACATGGACTTGCTCA | TACGTTCCGCAAACCTGTCA | pOX38-B |
| terC2 | TAAGGAGACCCTGTATGGC | CCACCAATCCGTTGATGTAA | <i>E. coli</i> chromosome |

Table SII: DNA fragments containing SopB-binding site used as *sopC* substrates.

| Names | SequencesFlanking DNA | size (bp) |
|------------------|---|-----------|
| <i>sopC</i> -144 | 43-bp..gaggtcgatgattattagtcTGGGACCACGGTCCCActcgtatcgtcgggtatcga.. 37-bp | 74 |
| <i>sopC</i> -60 | gaggtcgatgattattagtcTGGGACCACGGTCCCActcgtatcgtcgggtatcga | 22 |
| <i>sopC</i> -40 | tgattattagtcTGGGACCACGGTCCCActcgtatcgtcg | 12 |
| <i>sopC</i> -36 | attattagtcTGGGACCACGGTCCCActcgtatcgt | 10 |
| <i>sopC</i> -30 | attagtcTGGGACCACGGTCCCActcgtat | 7 |

The 16-bp SopB-binding site is indicated in upper case. Note that each DNA fragment is designated by its length in base pair (bp).

Table SIII: *E. coli* strains and plasmids

| Strains | Genotype/relevant properties | Source/references |
|----------|--|-------------------------|
| DLT1215 | <i>F⁻, thi, leu, thyA, deoB, supE, Δ(ara-leu)7696, zac3051::Tn10, rpsL812</i> | (Bouet et al. 2005) |
| DLT1472 | DLT1215 <i>lacZΩsopB⁺</i> | (Bouet et al. 2005) |
| DLT1780 | DLT1215/pDAG114 | This work |
| DLT2888 | DLT1215/pOX38-B conjugation from NCM3722/pOX38-B | This work |
| DLT2895 | Stellar/pJYB212 | This work |
| DLT2928 | Stellar/pJYB222 | This work |
| DLT2932 | Stellar/pJYB226 | This work |
| DLT2941 | Stellar/pJYB216 | This work |
| DLT2997 | Stellar/pJYB234 | This work |
| DLT3024 | Stellar/pJYB243 | This work |
| Stellar | <i>F⁻, λ⁻, endA1, supE44, thi-1, recA1, relA1, gyrA96, phoA, Δ80d lacZ ΔM15, Δ(lacZYA-argF) U169, Δ(mrr-hsdRMS-mcrBC), ΔmcrA</i> | Clontech (HST08) |
| Plasmids | Relevant characteristics | Source/references |
| pAM238 | <i>lacOPZ'Ωmcs</i> in <i>EcoRI-HindIII</i> interval of pGB2, <i>aadA⁺</i> | (Bouet et al. 1996) |
| pDAG114 | <i>repFIA⁺, ccdB⁻, resD⁺, rsfF⁺, cat⁺</i> | (Lemonnier et al. 2000) |
| pJYB212 | pDAG114, <i>sopB::mEos2</i> | This work |
| pJYB214 | pDAG114, <i>sopB::Venus-Yfp</i> | (Sanchez et al. 2013) |
| pJYB216 | pJYB212, <i>ΔsopC</i> | This work |
| pJYB222 | pJY212, <i>sopB-R219A::mEos2</i> | This work |
| pJYB226 | pJY212, <i>sopB-A183D-K191A-R195A::mEos2</i> | This work |
| pJYB234 | pDAG214, <i>sopB::mVenus-Yfp</i> | This work |
| pJYB243 | pDAG114, <i>sopA::mVenus-Yfp</i> | This work |
| pOX38-B | pOX38 <i>ΔccdB::cat</i> (<i>cat</i> was introduced in <i>ccdB</i> by lambda-red recombineering) gift from A. Wright | (Guyer et al. 1981) |
| pZC204 | pACYC184 <i>ΔtetAΔsopC⁺, cat⁺</i> | (Yates et al. 1999) |

Chapitre IV

Etudes préliminaires

I. Etat de l'Art

La protéine SopB, en plus de faire le lien physique entre le complexe de partition et le moteur de la partition SopA, est aussi capable de réguler les activités de SopA. En effet, SopB est importante pour que SopA oscille d'un pôle à l'autre du nucléoïde, et l'activité ATPase de SopA augmente en présence d'un ou des deux partenaires du complexe de partition (SopB/*sopC*) (Ah-Seng *et al.*, 2009). SopB régule également la polymérisation de SopA (Bouet, Ah-Seng, *et al.*, 2007). En effet, l'équipe a montré que la polymérisation spontanée en présence d'ATP était inhibée en présence d'ADN. Cependant, celle-ci est restaurée par la présence de SopB qui active directement la polymérisation de SopA et qui lève indirectement l'inhibition par l'ADN en masquant celui-ci (Bouet, Ah-Seng, *et al.*, 2007). Ainsi dans la cellule, la polymérisation de SopA serait inhibée dans le cytoplasme sauf à l'endroit où se trouve le complexe de partition "étendu", et où les protéines SopB se trouveraient concentrées. Ceci est à l'origine du modèle de "DNA free zone" (**Figure 33**), dans lequel SopB grâce à son activité d'étalement, pourrait créer une zone où la polymérisation pourrait avoir lieu.

Par ailleurs SopA est également capable de réguler les activités de SopB. A forte concentration, SopA est capable de déstabiliser le complexe de partition "étendu". Lors de l'oscillation, la rencontre de la vague de SopA avec les protéines SopB assemblées en complexe de partition "étendu", pourrait permettre la déstabilisation du complexe et à posteriori la séparation des plasmides.

L'interaction entre SopA et SopB assemblée au niveau du complexe de partition, est une étape clé qui pourrait contrôler la partition en régulant l'hydrolyse de l'ATP, la déstabilisation du complexe étendu et la polymérisation de SopA.

La protéine SopB est une protéine multi-fonctionnelle. La structure de la partie C-terminale de SopB (Schumacher *et al.*, 2010), révèle uniquement les domaines de dimérisation et de fixation à l'ADN. Mais comme pour les autres protéines formant un complexe de partition, la moitié N-terminale n'a pas été résolue, confirmant qu'elle serait flexible et pas fortement structurée (Schumacher and Funnell, 2005). Cette partie est pourtant importante car elle comporte les domaines (i) d'interaction avec l'ATPase, (ii) d'oligomérisation, et (iii) d'activation de l'activité ATPase de SopA (Surtees, 2001; Ravin *et al.*, 2003; Ah-Seng *et al.*, 2009). La détermination précise des domaines d'interaction présents dans la partie N-terminale de SopB apparaît donc essentielle.

Le but de l'étude qui suit est de caractériser précisément les domaines d'interaction protéine-protéine présents dans la partie N-terminale de SopB. La caractérisation du domaine d'interaction SopB-SopA nous permettra de voir si cette interaction est indispensable à la régulation des activités de SopA, tandis que la caractérisation du domaine d'oligomérisation nous précisera le rôle du complexe "étendu" dans la régulation de la polymérisation de SopA.

Dans un premier axe d'étude et par des approches de mutagenèse par délétion, les déterminants majeurs impliqués dans l'interaction de la protéine SopB avec SopA ont été identifiés par la technique de résonance plasmonique de surface (SPR). Nous avons ensuite montré que l'interaction entre ces deux protéines est un prérequis pour garder la capacité de SopB à stimuler l'activité ATPase et la polymérisation de SopA *in vitro*.

Dans un second axe, nous avons recherché des variants de SopB incapables de faire de l'oligomérisation avec différentes stratégies de mutagenèse. L'objectif était de discriminer les différents mutants par une approche *in vitro*, en utilisant leur capacité à former des complexes supplémentaires en gel d'électrophorèse (cf. **Résultats, Chapitre III**). La difficulté d'obtenir des protéines variantes purifiées nous a empêchés d'avoir des réponses tranchées sur la capacité de ces variants à réaliser ou non l'oligomérisation. Deux approches *in vivo* ont donc été envisagées : (i) microscopie à fluorescence et (ii) qChIP. Elles nous ont permis d'identifier cinq mutants dont la capacité d'oligomérisation semble sérieusement affectée.

Dans une dernière partie, nous nous sommes intéressés à la possibilité que SopB puisse s'étendre sur l'ADN chromosomique. Les données préliminaires obtenues en ChIP-seq semblent indiquer que SopB est capable de s'étendre en *cis* sur l'ADN plasmidique comme cela

a été vu dans le **chapitre III**, mais également en *trans* sur le nucléoïde. Ainsi, le nucléoïde pourrait soit contribuer à l'architecture globale du complexe de partition "étendu" et/ou être impliquer dans le processus de ségrégation.

II. Caractérisation du domaine d'interaction entre SopB et l'ATPase SopA

A. Approche de mutagenèse par délétion: Construction des 3 mutants, SopB Δ 23, SopB Δ 9 et SopB Δ 8

D'après des études réalisées par (Ravin *et al.*, 2003) ses collaborateurs, les 42 premiers résidus de SopB seraient impliqués dans la spécificité d'interaction avec SopA (Ravin *et al.*, 2003). En effet, l'échange de ces premiers acides aminés avec les résidus de son homologue fonctionnel ParB du prophage N15, ne permettent pas de préserver l'interaction spécifique de SopA avec SopB. Des analyses bio-informatiques de SopB et de ses homologues fonctionnels ont révélé que les 30 premiers acides aminés de ces protéines présentent une région hypervariable, suggérant que les déterminants de la spécificité de SopB pour SopA seraient présents dans cette région (Ah-Seng *et al.*, 2009) (**Figure 47a**). Afin de déterminer plus précisément la zone d'interaction de SopB avec SopA, nous nous sommes basés sur ces analyses de séquences pour construire des mutants de délétion de SopB. SopB Δ 23 correspond à une délétion de 23 aa englobant l'ensemble de la région hypervariable alors que SopB Δ 8 et SopB Δ 9, délétion de 8 et 9 aa respectivement, correspondent à 2 sous parties non chevauchantes de cette région (**Figure 47 b et c**). Ces délétions gardent l'arginine 36 impliquée dans la stimulation de l'activité ATPase de SopA, intacte.

B. Purification et implication de la partie N-terminale de SopB dans le comportement non globulaire de la protéine

Les protéines SopB et les 3 variants ont été purifiés en 3 étapes comme précédemment (cf. **Matériel et méthodes**) : une colonne d'affinité, une colonne de chromatographie d'exclusion suivie d'une colonne échangeuse de cations.

Le profil d'élution de la colonne de chromatographie d'exclusion, nous informe sur l'état oligomérique de la protéine ainsi que sur la conformation qu'elle adopte. Un profil d'élution

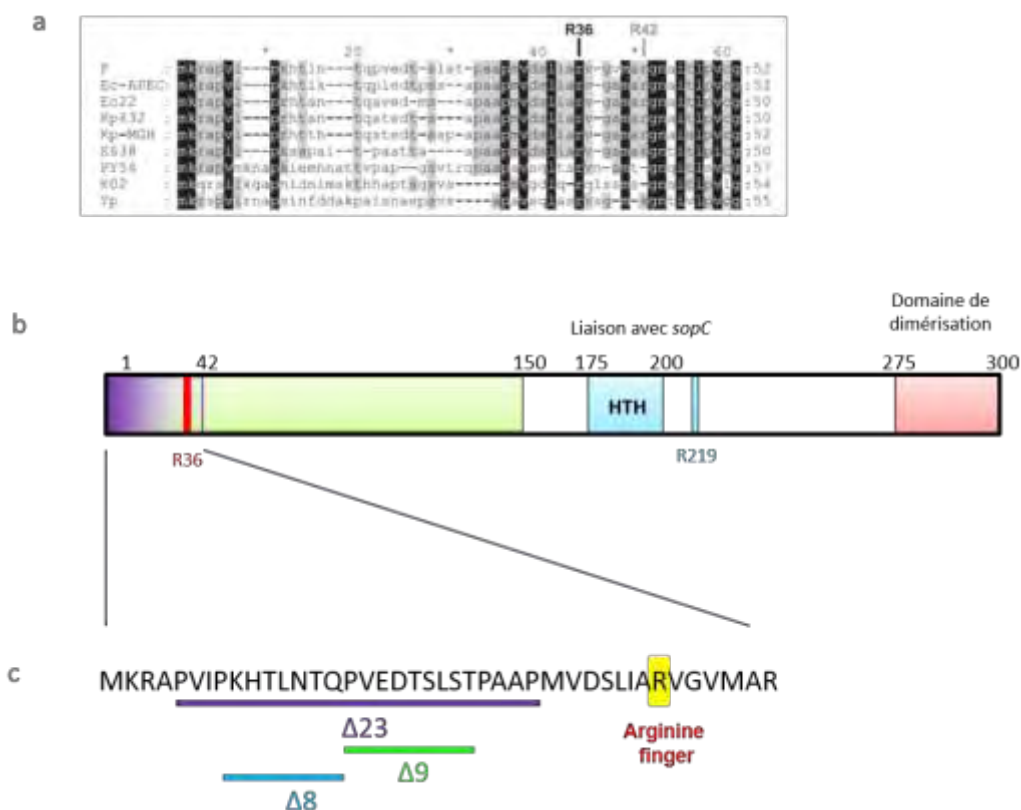


Figure 47 / Organisation fonctionnelle de SopB et carte des délétions construites

Les positions des aa de SopB sont indiquées au-dessus de la séquence (a) et de la représentation schématique (b)

- Alignement des séquences d'acides aminés de la région N-terminale des homologues de SopB. Des homologues très proches de SopB ont été trouvés dans la base de données en utilisant l'algorithme BLAST. Neuf d'entre eux, retrouvés sur des plasmides ou des phages, présentent une identité >30% au niveau de l'extrémité N-terminale (>45% dans toute la séquence) et ont également une arginine correspondant à l'arginine 36 de SopB. Dans les séquences alignées, les résidus conservés et les substitutions conservatives (V / I / L, T / S, et R / K) sont ombrés en noir si elles sont présentes dans toutes les séquences et en gris si elles sont présentes dans au moins cinq séquences. Les résidus arginine notés R36 et R42 se réfèrent à SopB de F. Les protéines ParB appartiennent aux plasmides d'*E. coli* APEC O1 (Ec-APEC), K12 (F), E22 (EC22), *Klebsiella pneumoniae* 342 (Kp342), *K. pneumoniae* MGH (Kp-MGH), *Enterobacter* sp. 638 (E638), *Yersinia pestis* bv Microtus (YP). Les protéines de phages homologues à SopB sont de *Yersinia* phage PY54 (PY54) et de *Klebsiella* phage phiKO2 (KO2).
- Domaines fonctionnels de SopB : SopB possède un motif Hélice tour hélice (HTH) en bleu, des résidus 179-200, responsable de la fixation avec le motif inversé-répété de *sopC*. Les résidus 275 à 323, en rouge, correspondent au domaine de dimérisation de SopB. Les résidus 1 à 150, en vert, pourraient être impliqués dans l'oligomérisation des dimères de SopB (Hanai *et al.*, 1996). La région des résidus 1 à 42, en mauve, pourrait être impliquée dans la spécificité d'interaction avec SopA (Ravin *et al.*, 2003)
- Carte des délétions effectuées dans SopB : Les mutants SopBΔ23, SopBΔ8, SopBΔ9 sont respectivement délétés de 23 aa (position [5 à 27]), 8 aa (position [8 à 15]) et 9 aa (position [16 à 24]). L'arginine 36 est l'arginine finger de SopB, responsable de la stimulation de l'activité ATPase de SopA, surlignée en jaune.

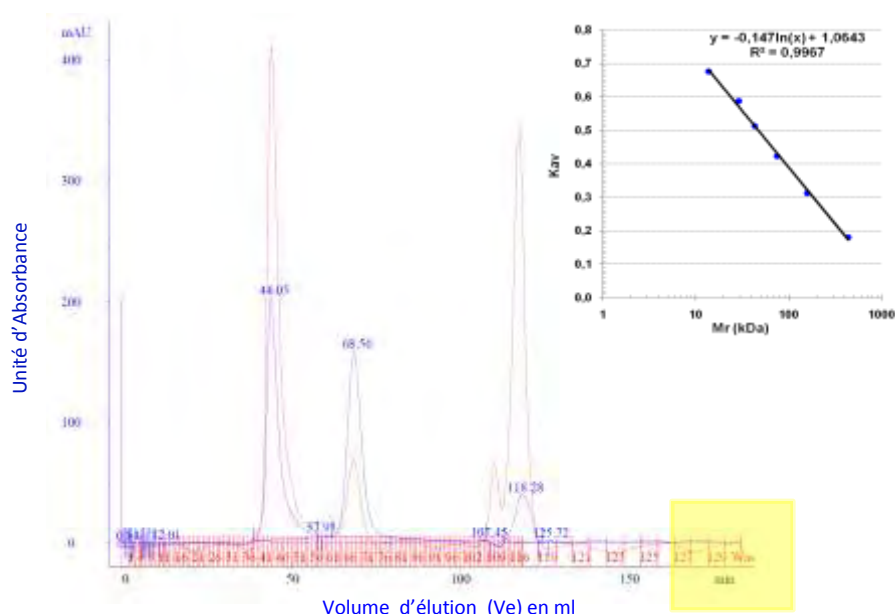


Figure 48 Profil d'élution de SopBΔ8 sur une colonne de chromatographie d'exclusion

La fraction II concentrée est séparée sur une colonne de chromatographie d'exclusion (superdex 200). La courbe en rouge correspond à l'absorption à 260 nm (longueur d'onde d'absorption des acides nucléiques) et celle en bleu à l'absorption à 280 nm (longueur d'onde d'absorption des protéines). Le profil d'élution est formé de 3 pics principaux. Le premier pic à 44.05 ml, représentant le volume mort de la colonne. Il pourrait correspondre à des acides nucléiques et/ ou des protéines agrégées. Le deuxième pic à 68.56 ml, représente le volume d'élution de SopBΔ8. Le troisième pic est le volume de la colonne et correspond à l'élution des petites molécules. Ici, ce pic pourrait être du DTT qui a été rajoutée dans le tampon.

Encart : Calibration de la colonne de la superdex 200. La droite présentée dans ce graphique correspond à l'élution de 6 protéines de poids moléculaire connu et qui possèdent un comportement globulaire. Les valeurs sont calculées par la formule suivante : $K_{av} = V_e - V_o / V_c - V_o$ (K_{av} : coefficient de distribution dans la colonne, V_e : volume d'élution, V_o : volume mort = 39.8 ml et V_c : volume géométrique de la colonne = 121ml). Ce graphique nous permet de calculer le poids moléculaire apparent (Tableau 5), des protéines SopB, SopBΔ23, SopBΔ9 et SopBΔ8.

| Protéines | Volume d'élution (ml) | MM apparente (kDa) | MM d'un monomère | MM d'un dimère théorique (kDa) | MM d'un tétramère théorique (kDa) | ratio MM apparente/ monomère théorique |
|-----------|-----------------------|--------------------|------------------|--------------------------------|-----------------------------------|--|
| SopB | 67,6 | 136,7 | 35,37 | 71 | 141,5 | 3,9 |
| SopBΔ23 | 71,4 | 99,1 | 32,97 | 66 | 131,9 | 3 |
| SopBΔ9 | 69,05 | 120,9 | 34,44 | 68,9 | 137,8 | 3,5 |
| SopBΔ8 | 68,58 | 120,9 | 34,45 | 68,9 | 137,8 | 3,5 |

MM, Masse Moléculaire

Tableau 5 Analyse conformationnelle de la partie N-terminale de SopB et de ces variants en fonction de leur comportement en chromatographie d'exclusion.

Le volume d'élution correspond au volume d'élution de la protéine en chromatographie d'exclusion. La masse moléculaire (MM) apparente est le MM déduite de la courbe de calibration de la colonne de chromatographie d'exclusion, étalonnée avec des protéines globulaires. Les MM théoriques du dimère et du tétramère sont les MM calculées à partir de la MM théorique du monomère. Le ratio MM apparente/ monomère théorique traduit le nombre de monomère possible par rapport à la MM apparente.

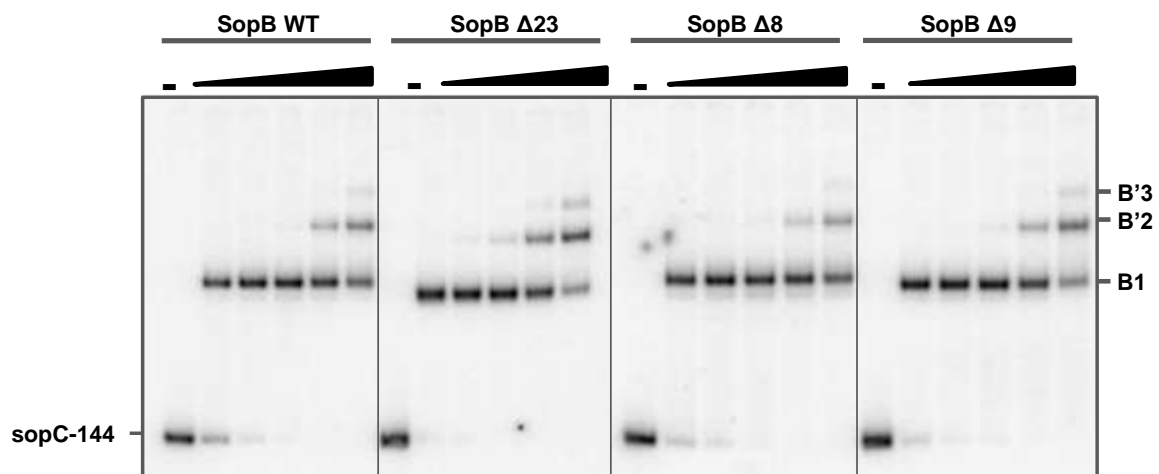


Figure 49 | *Activité de fixation spécifique de SopB et des variants SopBΔ23, SopBΔ9 et SopBΔ8 sur sopC*

La fixation de SopB à *sopC* est visualisée par des expériences de retard sur gel. La sonde spécifique (*sopC*), marquée au ^{32}P , est incubée seule (piste 1) ou avec une augmentation croissante de protéine SopB, SopBΔ23, SopBΔ9 et SopBΔ8 (piste 2-5 ; 0, 10, 30, 100, 300 et 1000 nM) en présence de $100\mu\text{g.ml}^{-1}$ d'ADN compétiteur. Les mélanges réactionnels sont analysés par électrophorèse sur gel d'acrylamide. La position de la sonde et des complexes (B1), (B'2), (B'3) et de la sonde libre est indiquée sur la gauche.

type est montré en **Figure 48**. La calibration de cette colonne (encart de la **Figure 48**) a été préalablement faite en établissant le volume nécessaire à l'élution de différentes protéines globulaires de masse moléculaire (MM) connue. Ceci nous a permis de déduire la MM apparente des différentes protéines SopB purifiées. Les volumes d'élution et les MM apparentes déduites de SopB et des variants SopB Δ 23, SopB Δ 9 et SopB Δ 8 sont montrées en **Tableau 5**. D'après les données de la littérature, SopB est un dimère en solution et se fixe spécifiquement sur *sopC* sous forme dimérique (Hanai *et al.*, 1996). Sa MM attendue sous forme dimérique est de 71 kDa alors que sa MM apparente est de 136.7 kDa. De même, les MM apparentes des variants SopB Δ 23 (99 kDa), SopB Δ 9 (120.9 kDa) et SopB Δ 8 (120.9 kDa), ne correspondent pas non plus à des protéines dimériques.

Le calcul du ratio, MM apparente sur monomère est de 3.9 pour SopB et de 3 pour SopB Δ 23. Les variants SopB Δ 9 et SopB Δ 8 ont un ratio intermédiaire compris entre 3 et 4. Le comportement des protéines en chromatographie d'exclusion n'est pas seulement lié au nombre de sous-unités associées en solution mais dépend également de la conformation tridimensionnelle de ces protéines. Le changement de comportement entre SopB et SopB Δ 23 indique que les 23 aa délétés sont responsables de 50% du comportement non globulaire de la protéine.

L'hypothèse la plus probable est que contrairement aux protéines utilisées pour la calibration, SopB n'est pas une protéine globulaire mais constituée de régions non structurées, ce qui est en accord avec les données structurales et biochimiques de la littérature. En effet, la région de SopB (1-150) ou de son homologue ParB du plasmide P1 semble être flexible, comme en témoigne également, la sensibilité de cette région à la protéolyse (Surtees and Funnell, 1999) et la difficulté à obtenir des cristaux analysables (Schumacher and Funnell, 2005; Schumacher *et al.*, 2010).

Les 23 aa ciblés pour réaliser les délétions de SopB et susceptibles d'être impliqués dans l'interaction avec SopA, semblent faire partie d'une région non structurée.

C. Interaction de SopB et de ces variants avec le centromère, *sopC*

La fonctionnalité des protéines a été testée pour leur capacité à fixer de façon spécifique *sopC* par des expériences de retard sur gel. Pour cela, nous avons utilisé comme substrat, une sonde radiomarquée (*sopC144*), possédant un seul motif consensus de fixation

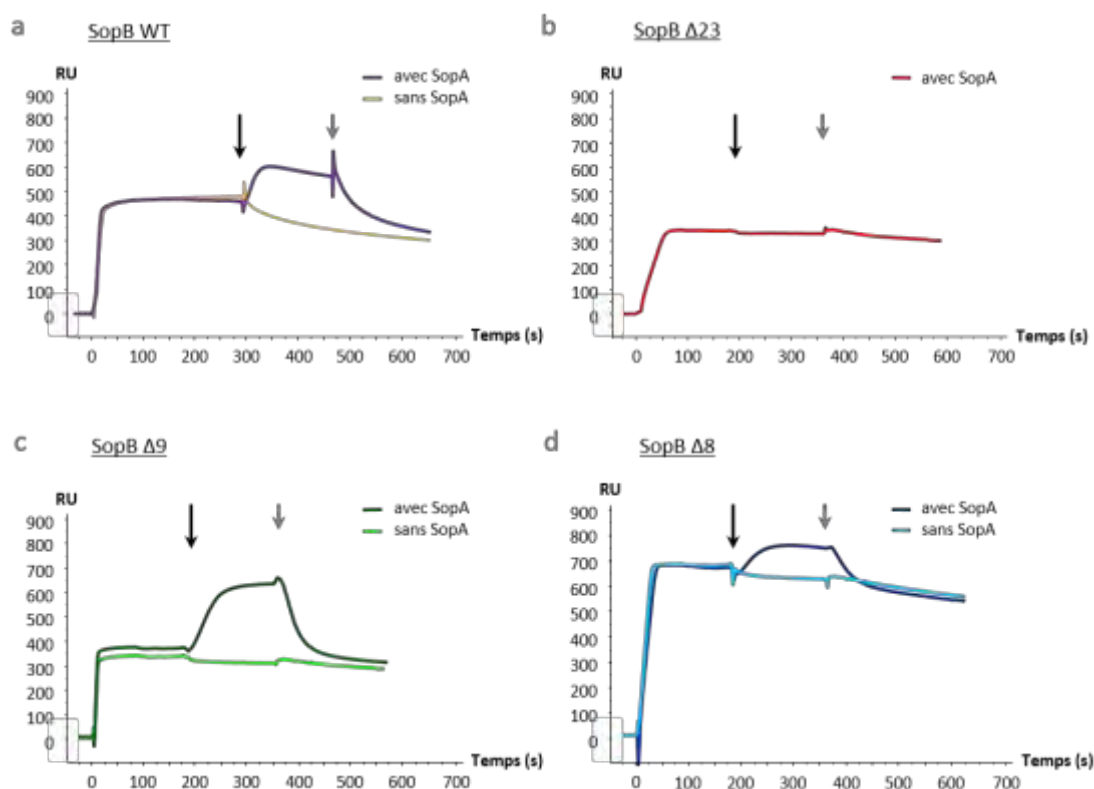


Figure 50 *Analyse en SPR de la fixation de SopA sur SopB ou ces variants SopBΔ23, SopBΔ9 et SopBΔ8 immobilisés sur des sondes sopC*

Les résultats sont exprimés comme la différence de RU (Unité de Réflectivité) entre le signal obtenu avec une sonde spécifique (*sopC*) de 136 pb et une sonde non spécifique de 136 pb en fonction du temps.

(a) Liaison de SopA avec le complexe SopB/*sopC*. Le tampon contenant SopB à 150 nM, est injecté dans le tampon sur les sondes ADN immobilisées, à temps zéro. Du tampon seul ou SopA (500 nM) en présence d'ATP, est ensuite injectée dans le flux de tampon (flèche noire). Les complexes sont lavés avec du tampon sans protéine au temps indiqué par la flèche grise. (b-d) comme en (a) à l'exception des protéines SopBΔ23, SopBΔ9, SopBΔ8 qui sont injectées respectivement à 200nM, 200nM et 400nM. Pour SopB-Δ23 (b), seule la courbe avec SopA a été obtenue.

de SopB. Les expériences réalisées en présence d'ADN compétiteur pour visualiser le retard spécifique, sont effectuées avec une gamme croissante de SopB ou des variants de SopB (**Figure 49**). Pour SopB, la sonde *sopC144* est retardée à partir d'une concentration de 10nM de SopB.

Ce retard (B1) est discret et proportionnel à la concentration en SopB apportée. On observe un 2^{ème} et un 3^{ème} retard (B'2 et B'3) de la sonde spécifique, reflétant la capacité de SopB à s'oligomériser et à s'étaler de manière non spécifique sur l'ADN à partir du point de nucléation *sopC* (cf. **Résultats, Chapitre III**). De la même manière SopB Δ 23, SopB Δ 8 et SopB Δ 9 gardent leur capacité à fixer spécifiquement *sopC* et se fixer sur l'ADN flanquant à partir du point de nucléation *sopC* (**Figure 49**). Les 3 protéines tronquées sont donc fonctionnelles pour leur interaction avec *sopC*.

D. Mesure d'interaction de SopA avec SopB ou les protéines SopB tronquées

L'étude des interactions moléculaires en temps réel, utilisant la résonance plasmonique de surface (SPR) a déjà permis de visualiser l'interaction de SopA, avec SopB liée à *sopC*, en présence d'ATP (Ah-Seng *et al.*, 2009). La stratégie que nous avons adoptée pour caractériser le domaine d'interaction de SopB avec SopA, a consisté à tester l'interaction des variants, SopB Δ 23, SopB Δ 9 et SopB Δ 8 avec SopA par la même technique d'analyse.

Dans cette expérience, de l'ADN contenant un site de fixation de SopB (*sopC*) ou de l'ADN non spécifique (ADN-NS), est préalablement immobilisé à la surface du support. La quantité d'ADN immobilisé est proportionnelle à la variation de l'indice de réfraction, donnée en unité de résonance (RU). La fixation spécifique de SopB sur la sonde *sopC* est visualisée par une courbe différentielle entre la piste contenant la sonde *sopC* et celle contenant l'ADN-NS.

Après injection de SopB, la fixation spécifique de SopB à *sopC* se traduit par l'augmentation du signal SPR (**Figure 50a**). Cette augmentation est aussi visualisable pour les variants de SopB (**Figure 50b-d**). Ainsi, SopB Δ 23, SopB Δ 9, SopB Δ 8 ont gardé la capacité de fixer *sopC*, confirmant les résultats obtenus par les expériences de retard sur gel. Dans cette expérience, SopB et ses variants sont injectés pour permettre de saturer la sonde *sopC*. La quantité d'ADN déposée est différente suivant les expériences mais toujours quantifiable (proportionnelle à l'indice de réfraction, RU). La saturation des sites *sopC*, par les protéines

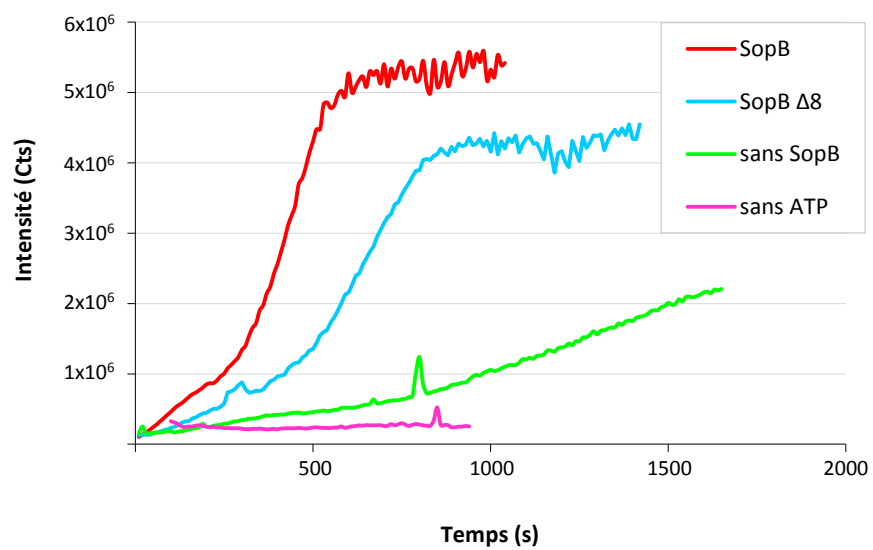


Figure 51 *Effet des 3 variants de SopB sur la stimulation de la polymérisation de SopA*

La capacité de polymérisation de SopA est mesurée en dynamic light scattering (DLS). Toutes les expériences contiennent SopA (4 μ M) et 1mM d'ATP sauf la courbe marron (sans ATP). Les protéines SopB (courbe rouge), SopB $\Delta 9$ (courbe verte) et SopB $\Delta 8$ (courbe bleu) sont ajoutées à 2 μ M. La réaction est initiée à un temps zéro par l'ajout d'ATP. L'intensité de polymérisation, mesurée en coups (cts), est suivie en fonction du temps.

SopB a ainsi été obtenue avec 470 RU pour SopB, 640 RU pour SopB Δ 8 et 350 RU pour SopB Δ 23 et SopB Δ 9.

Après formation du complexe SopB/*sopC*, SopA-ATP est injectée dans le flux. Une augmentation du signal SPR (de l'ordre de 150 RU) est détectée, indiquant une interaction entre SopA-ATP et SopB liée à *sopC* (**Figure 50**). Quand SopA est injectée après formation de SopB Δ 23/*sopC* (**Figure 50b**), aucune augmentation dans le signal SPR n'a été détectée, indiquant que le variant SopB Δ 23 a entièrement perdu la capacité d'interagir avec SopA. Pour le variant SopB Δ 9, on observe une interaction avec SopA (200 RU) comparable à celle obtenue avec SopB pleine taille, indiquant que l'interaction avec SopA n'a pas été significativement affectée par cette délétion. Dans l'expérience avec SopB Δ 8, on a déposé 2 fois plus de sonde que pour SopB Δ 9 (640 RU contre 350 RU). La réponse SPR (100 RU) suite à la fixation de SopA-ATP sur le complexe SopB Δ 8/*sopC* est deux fois moins importante que dans l'expérience avec SopB Δ 9. La diminution du signal est de l'ordre d'un facteur 4, indiquant que le variant SopB Δ 8 est très affecté pour l'interaction avec SopA.

Le variant SopB Δ 23 a totalement perdu sa capacité d'interaction avec SopA, suggérant que le domaine d'interaction de SopB avec SopA est compris entre les résidus 5 à 28 de SopB. Le variant SopB Δ 9 (délétion des résidus de 16 à 24) est toujours capable d'interagir avec SopA tandis que le variant SopB Δ 8 est très affecté pour l'interaction. L'ensemble de ces données suggère que les déterminants d'interactions majeurs de SopB avec SopA, pourraient se situer dans les 15 premiers aa.

E. Effet des variants de SopB sur la stimulation de la polymérisation de SopA

Des expériences réalisées par des analyses de dynamic light scattering (DLS) ont montré que SopA a la capacité de polymériser *in vitro*, en présence d'ATP. De plus, cette polymérisation est stimulée par SopB (Bouet, Ah-Seng, *et al.*, 2007). La capacité de stimulation de la polymérisation de SopA par SopB et des variants a été mesurée en DLS afin de déterminer le rôle de leur interaction dans la stimulation de la polymérisation de SopA. Cette technique utilise la diffusion de la lumière dynamique pour déterminer le profil de distribution de la taille des molécules en solution. L'expérience se déroulant à concentration constante en protéine, le profil de distribution traduit le degré de polymérisation de SopA dans la solution.

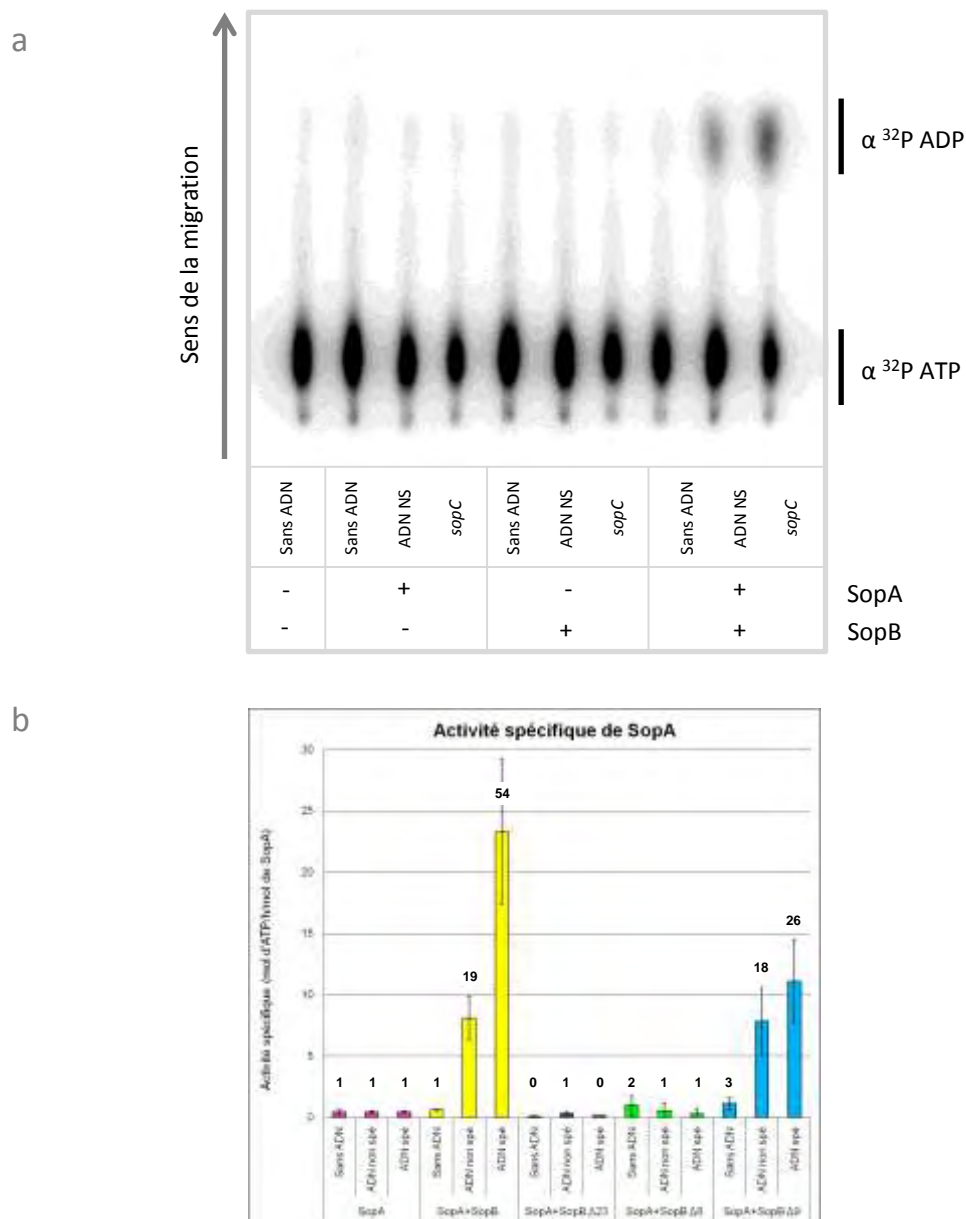


Figure 52 *Activation de l'activité ATPase de SopA par SopB ou ces variants SopBΔ23, SopBΔ9 et SopBΔ8*

Les mélanges réactionnels contiennent 2 μM de SopA et 2 μM de SopB ou de SopBΔ23, SopBΔ9 ou SopBΔ8, 60 μg.ml⁻¹ d'ADN non spécifique (pBSKS) ou d'ADN contenant *sopC* (pJYB57) sont aussi ajoutés au mélange ou non comme indiqué.

- a) Chromatographie sur plaque de TLC, obtenue après dépôts, migration des mélanges réactionnels puis révélation par autoradiographie. La migration de l' α ³²P ADP et de l' α ³²P ATP est indiquée sur la droite. La présence ou non de SopA, SopB et de l'ADN non spécifique (ADN NS) ou d'un ADN contenant *sopC* dans le mélange est indiquée sous la figure.
- b) Histogramme représentant l'activité ATPase, spécifique, moyenne de SopA, mesurée en molécule d'ATP hydrolysée par heure par molécule de SopA. Pour chaque valeur, la valeur correspondant au contrôle sans SopA est soustraite. Les moyennes des activités spécifiques sont réalisées suite à au moins 3 expériences indépendantes. Les facteurs d'activation de l'activité d'hydrolyse de SopA sont marqués en haut de chaque histogramme.

La polymérisation de SopA est strictement dépendante de l'ATP (comparer courbe rose et courbe verte) (**Figure 51**). La polymérisation est stimulée par la présence de SopB (courbe rouge) ce qui correspond aux données obtenues précédemment (Bouet, Ah-Seng, *et al.*, 2007). Au cours du temps la courbe de polymérisation atteint un plateau, reflétant un état d'équilibre entre polymérisation et dépolymérisation de SopA. Pour des raisons techniques, sur les 3 mutants étudiés, seul le variant SopB Δ 8 a pu être analysé en DLS. Ce variant stimule moins efficacement la polymérisation de SopA (courbe bleu). On observe une polymérisation plus lente et moins importante que pour SopB. Ce résultat est donc corrélé avec la perte partielle d'interaction de ce variant avec SopA, observée en SPR.

Ces données suggèrent que la capacité de stimulation de la polymérisation de SopA par SopB est étroitement liée avec sa capacité d'interaction avec SopA et devra être confirmée par l'étude du variant SopB Δ 23, incapable d'interagir avec SopA.

F. Effet des variants de SopB sur la stimulation de l'activité ATPase de SopA

L'activité ATPase de SopA (0.3 mol d'ATP/h/mol de SopA) est faiblement activée par l'ADN (environ 2 fois) et par SopB (environ 3 fois) mais elle est fortement activée (environ 40 fois) par la présence simultanée d'ADN non spécifique et de SopB (Bouet, Ah-Seng, *et al.*, 2007). La présence simultanée de SopB et d'ADN spécifique contenant *sopC* augmente de 120 fois la capacité d'hydrolyse de SopA (36 mol ATP/h/mol SopA) (Ah-Seng *et al.*, 2009). Ceci représente une stimulation supplémentaire d'un facteur 3 par rapport à l'ADN non spécifique, suggérant que la présence du complexe de partition active l'hydrolyse de l'ATP par SopA. Par ailleurs les auteurs ont aussi montré que l'arginine 36 de SopB, est une arginine finger nécessaire à la stimulation de l'activité ATPase de SopA. Dans la première partie de l'étude, j'ai pu montrer que les résidus 5 à 15 pourraient être impliqués dans l'interaction SopB/SopA. Cette interaction est-elle une condition nécessaire à l'activation de l'activité ATPase de SopA par SopB ?

Les tests d'activité ATPase de SopA se font en présence ou en absence d'ADN spécifique contenant un motif *sopC*, d'ADN non spécifique et de SopB. SopA et SopB sont apportées en concentration équimolaire (2 μ M). La **Figure 52a** montre les résultats typiques d'une chromatographie en couche mince, permettant de visualiser le niveau d'hydrolyse.

Dans la série d'expériences réalisée pour cette étude, la capacité d'hydrolyse de SopA est augmentée de 19 fois en présence de SopB et d'ADN non spécifique et de 54 fois en

présence de SopB et d'ADN spécifique. Pour le variant SopB Δ 23, incapable d'interagir avec SopA, on observe une perte totale de la capacité de stimulation de l'hydrolyse de l'ATP dans toutes les conditions. En ce qui concerne SopB Δ 8, on observe une perte totale de la capacité de stimulation de l'hydrolyse de l'ATP alors qu'il n'avait qu'une perte partielle d'interaction avec SopA.

La stimulation de l'hydrolyse de SopA en présence simultanée de SopB Δ 9 et d'ADN non spécifique ne semble pas affectée puisqu'elle est augmentée de 18 fois comme le sauvage. Par contre, la présence d'ADN spécifique ne stimule plus que d'un facteur 1.4 supplémentaire au lieu de 2.9 (**Figure 52b**). L'une des hypothèses est que la fixation de SopB sur *sopC* induit un changement conformationnel de SopB, permettant une meilleure activation par l'arginine finger (R36) de SopB. Ainsi, SopB, pourrait moduler en fonction du type d'ADN fixé, la stimulation de l'activité ATPase. Dans le cas du mutant SopB Δ 9, la délétion pourrait affecter le changement de conformation de SopB dépendant de *sopC* et se traduire par la perte de cette sur-stimulation de l'activité ATPase de SopA.

L'ensemble des résultats suggère que l'activité ATPase de SopA pourrait nécessiter une interaction forte entre SopB et SopA. La perte d'une partie des déterminants d'interaction entre SopB et SopA, pourrait suffire à perdre totalement la capacité de stimulation de l'activité ATPase de SopA par SopB.

G. Conclusion-Discussion

Cette étude indique que le domaine d'interaction de SopB avec SopA pourrait se situer entre les résidus 5 à 15 de SopB. Cependant il ne faut pas exclure que le domaine d'interaction soit plus étendu avec des résidus non essentiels en dehors de ce domaine. La recherche des acides aminés impliqués dans cette interaction pourra être envisagée ultérieurement. En plus de définir le domaine d'interaction de SopB avec SopA, cette étude nous a fourni des mutants d'interaction qui sont de véritables outils pour la compréhension des mécanismes moléculaires impliqués dans la régulation des activités de SopA ainsi que dans le mécanisme global de partition.

Ainsi, nous avons étudié l'influence de l'interaction de SopA avec SopB sur deux activités de SopA, que sont l'activité ATPase et l'activité de polymérisation. Notre étude a montré que la stimulation de ces activités par SopB est étroitement liée à l'interaction de ces deux protéines.

En effet, SopB Δ 23 et SopB Δ 8 n'interagissant pas correctement avec SopA, ne permettent plus de stimuler à la fois l'ATPase et la polymérisation de SopA contrairement à SopB Δ 9 dont l'interaction avec SopA est préservée. Cependant, il semble que la régulation de l'activité ATPase de SopA par SopB soit beaucoup plus stringente que pour la polymérisation de SopA. En effet, dans le cas de SopB Δ 8, la perte d'une partie des déterminants importants pour l'interaction avec SopA, se traduit par une perte totale de la stimulation de l'activité d'ATPase de SopA. Une stimulation optimale de cette activité pourrait nécessiter une interaction stable entre SopA et SopB permettant le rapprochement entre l'arginine finger de SopB et la poche catalytique de SopA.

Au cours de cette étude, nous avons testé l'effet de la perte d'interaction SopB/SopA sur les activités de SopA en utilisant exclusivement des approches *in vitro*. Afin de compléter l'influence de cette perte d'interaction sur la partition du plasmide F, des approches *in vivo* devront également être réalisées. Il serait donc intéressant de savoir si la perte d'interaction de SopB avec SopA a une importance dans la partition. Pour cela, il pourrait être envisagé de réaliser des tests de stabilité et de localisation du plasmide mini-F, dans une souche bactérienne exprimant les mutants d'interaction.

L'interaction entre SopB et SopA pourrait également avoir une influence sur la capacité de SopA à osciller d'un pôle à l'autre du nucléoïde. En effet, SopB est capable de stimuler les activités de SopA, dont l'activité ATPase. En absence de SopB, SopA n'est plus capable d'osciller. La première idée était de penser que la stimulation de l'hydrolyse de l'ATP de SopA par SopB, était nécessaire à l'oscillation. Dans une cellule contenant le mutant SopB-R36A, incapable de stimuler l'activité ATPase de SopA (SopB-R36A), SopA est toujours capable d'osciller, suggérant que ce n'est pas la stimulation de l'hydrolyse de l'ATP de SopA par SopB qui est à l'origine de l'oscillation (Ah-Seng *et al.*, 2013). Par ailleurs, la présence de SopB est un prérequis pour que SopA oscille. Ainsi, l'effet de SopB sur la capacité de SopA à faire l'oscillation, est peut être une conséquence de leur interaction. SopA comme ParA du plasmide P1 possède différentes conformations étant chacune impliquée à différents niveaux dans le mécanisme global. On ne connaît pas pour SopA, les détails de ces changements de conformation, cependant l'interaction entre SopB et SopA pourrait induire le changement de conformation nécessaire à l'oscillation. Une étude du mutant d'interaction entre SopB et SopA réalisé dans cette étude pourrait y répondre.

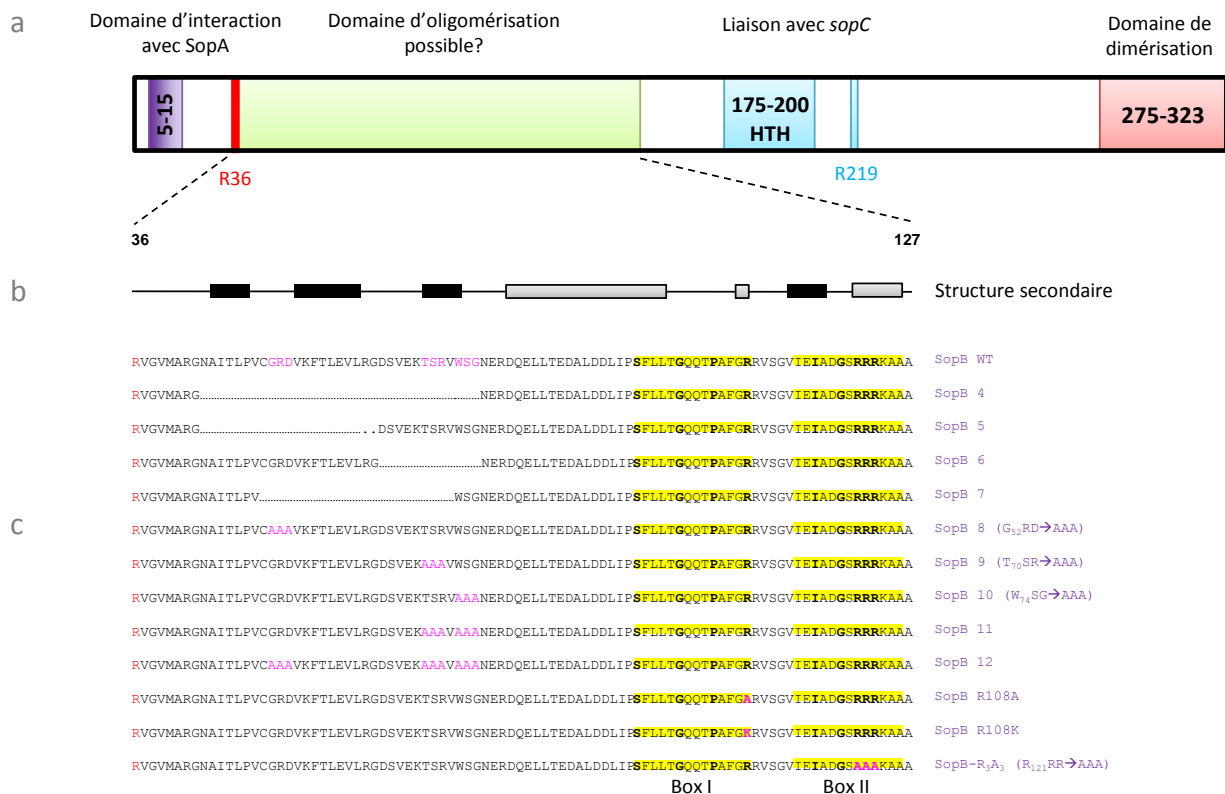


Figure 53 | Organisation fonctionnelle de SopB et carte des délétions construites

- a) Domaines fonctionnels de SopB :** SopB possède un domaine de fixation à *sopC*, composé d'un motif Hélice tour hélice (HTH) en bleu, des résidus 179-200, et de l'arginine 219 (R219). Les résidus 275 à 323, en rouge, correspondent au domaine de dimérisation de SopB. Les résidus 36 à 150, en vert, pourraient être impliqués dans l'oligomérisation des dimères de SopB. La région des résidus 5 à 15, en mauve, semble contenir les déterminants de la spécificité d'interaction avec SopA. J). L'arginine 36, surlignée, en jaune, est l'arginine finger de SopB, responsable de la stimulation de l'activité ATPase de SopA.
- b) Structure secondaire des 137 premiers aa de SopB,** prédite, par le logiciel "consensus secondary structure prediction" du site Pôle Bio-informatique Lyonnais (<http://npsa-pbil.ibcp.fr>). Les feuillets β sont représentés par des rectangles noirs et les hélices α par des rectangles gris.
- c) Carte des délétions et mutations effectuées dans SopB:** Les positions des aa de SopB sont indiquées au-dessus de la séquence et de la représentation schématique. Les délétions et les positions mutées sont indiquées en rose sur chaque séquence. Le nom des mutants est indiqué à gauche des séquences correspondantes. Les box I et II déterminées par l'alignement de séquences de différents homologues de SopB sont surlignées en jaune (Yamaichi & Niki, 2000). Les aa les plus conservés de ces boîtes sont indiqués en gras.

III. Caractérisation du domaine d'oligomérisation de SopB

La recherche de mutant d'oligomérisation des protéines de type ParB plasmidique ou chromosomique a fait l'objet de nombreuses études. Cependant, SopB étant une protéine multifonctionnelle, les études se sont heurtées à la difficulté d'obtenir un mutant qui ne soit déficient que pour cette fonction.

A. Approche de mutagenèse par délétion et mutation ponctuelle

Des études antérieures sur SopB (Hanai *et al.*, 1996) et sur son homologue ParB du plasmide P1 (Surtees, 2001) indiquent que les déterminants importants pour l'oligomérisation de ces protéines résident dans le domaine N-terminal, et que ce domaine pourrait être chevauchant avec le domaine d'interaction avec SopA/ParA. L'étude de différents mutants de SopB, incapables de faire du "Silencing" (éteindre des gènes à proximité de *sopC*), a permis de préciser que ce domaine d'oligomérisation pouvait être localisé entre les résidus 36 et 72 (Hanai *et al.*, 1996). D'après l'étude de la section précédente, ce domaine se situe hors des déterminants majeurs d'interaction avec SopA (5-15). Ainsi, nous nous sommes basés sur ces analyses pour construire quatre mutants de délétions. Puis, pour éviter de déstructurer la protéine, nous avons pris en compte la prédiction de la structure secondaire de SopB, afin de réaliser les délétions au niveau des régions non structurées et en évitant de couper au niveau d'une hélice α ou d'un feuillet β (**Figure 53**). Des études préliminaires réalisées dans l'équipe, avaient montré que trois mutants, SopB-8, SopB-9 et SopB-10, situés dans cette même région étaient incapables de faire du "Silencing". Ces mutants ont aussi été intégrés dans l'analyse. En combinant les mutants SopB-9 et 10 et SopB-8, 9 et 10, nous avons obtenu les mutants SopB-11 et 12 respectivement. Enfin, le motif B de SopB, motif conservé et constitué de deux boîtes Box I et Box II, pourrait également être impliqué dans l'oligomérisation. L'arginine 108 et les arginines 121, 122 et 123 sont très conservées chez tous les homologues de SopB (Yamaichi and Niki, 2000). Bien que ce motif soit situé en dehors de la région définie par Hanai et ses collaborateurs, trois mutants précédemment construits dans l'équipe, SopB-R₁₀₈K et SopB-R₁₀₈A situés dans la Box I et SopB-R₃A₃ (SopB-R₁₂₁A-R₁₂₂A-R₁₂₃A) dans la Box II, ont aussi été intégrés dans l'analyse. Ces douze mutants précédemment cités ont été intégrés par clonage dans le gène *sopB-intéine* sur le vecteur d'expression pYAS6 puis purifiés selon le principe précédemment décrit. Les protéines SopB-8 et 10 ne sont pas stables durant la

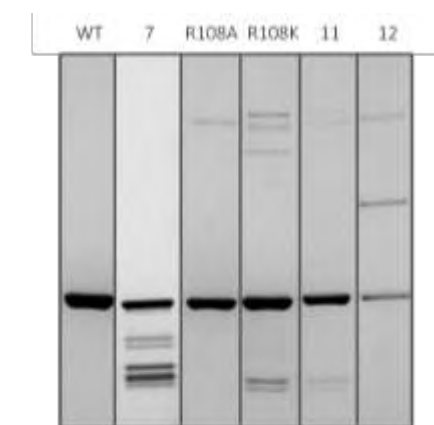


Figure 54 | Dernière étape de purification des protéines SopB-WT et des variants

Toutes les protéines ont été purifiées selon le protocole décrit dans le matériel et méthode. Un aliquot de la dernière étape de purification est mis à migrer sur gel SDS-PAGE pour chaque protéine. Puis les protéines SopB-WT, SopB-7, SopB-R₁₀₈A, SopB-R₁₀₈K SopB-11 et SopB-12 sont visualisées après coloration au bleu de Commassie. Dans la première piste, on note que la protéine SopB-WT purifiée est >95%. Les variants de SopB-WT contiennent des contaminants. Le profil de bandes contaminantes du variant SopB-7 est caractéristique des mutants de délétions comme SopB-4, SopB-5 et SopB-6, donc seul ce mutant est représenté.

purification. Par conséquent, elles ont été écartées de l'analyse. La pureté des protéines est vérifiée sur gel de Coomassie et montrée en **Figure 54**.

B. Profil de fixation sur un ADN contenant le site centromérique *sopC*

La fonctionnalité des protéines a été testée pour leur capacité à fixer de façon spécifique *sopC*, par des expériences de retard de migration sur gel. Les expériences sont réalisées comme précédemment (cf. **Résultat, Chapitre IV.II.C**), en présence d'ADN compétiteur pour visualiser les retards spécifiques (**Figure 54**). En présence de SopB, la sonde *sopC-144* est retardée de manière spécifique à partir d'une concentration de 10nM de SopB. On observe un 2^{ème} et un 3^{ème} retard (B'2 et B'3) de la sonde, reflétant la capacité de SopB à s'oligomériser et à se fixer de manière non spécifique sur l'ADN flanquant à partir du point de nucléation *sopC*. Tous les variants de SopB sont capables de retarder la sonde spécifique donc ils ont gardé leur capacité à se fixer spécifiquement sur le centromère.

Tous les mutants sauf SopB-R₁₀₈A et SopB-R₃A₃, ont des profils de migration qui ne correspondent pas à SopB. On observe des bandes multiples de différentes tailles au lieu d'avoir un complexe B1, B'2 et B'3 comme pour SopB. Ces bandes pourraient correspondre aux protéines de masse moléculaire inférieure à SopB qui étaient présent dans les gels SDS-Page (**Figure 55**). L'explication la plus probable est que toutes les mutations réalisées, semblent rendre instable la protéine et se traduit par l'apparition de produit de dégradation de différentes tailles. L'une des étapes de purification étant une colonne d'affinité avec une étiquette intéine située en C-terminal, la dégradation des différents mutants doit se produire du côté N-terminal. Plusieurs méthodes de purification alternatives ont été testées pour pallier à ces produits de dégradations secondaires : (i) utilisation d'une anti-protéase lors de l'étape de resuspension du culot (Frl), (ii) production des différents variants de SopB dans une souche BL21 délétée pour l'une des principales protéases, ClpP (iii) purification de la protéine native sans étiquette et avec un protocole différent. Aucune de ces méthodes n'a résolu les problèmes de dégradation. L'une des solutions de purification qui pourra être envisagée est de garder l'étiquette intéine labile en C-terminale et de rajouter une étiquette en N-terminale afin de purifier la forme pleine taille des protéines.

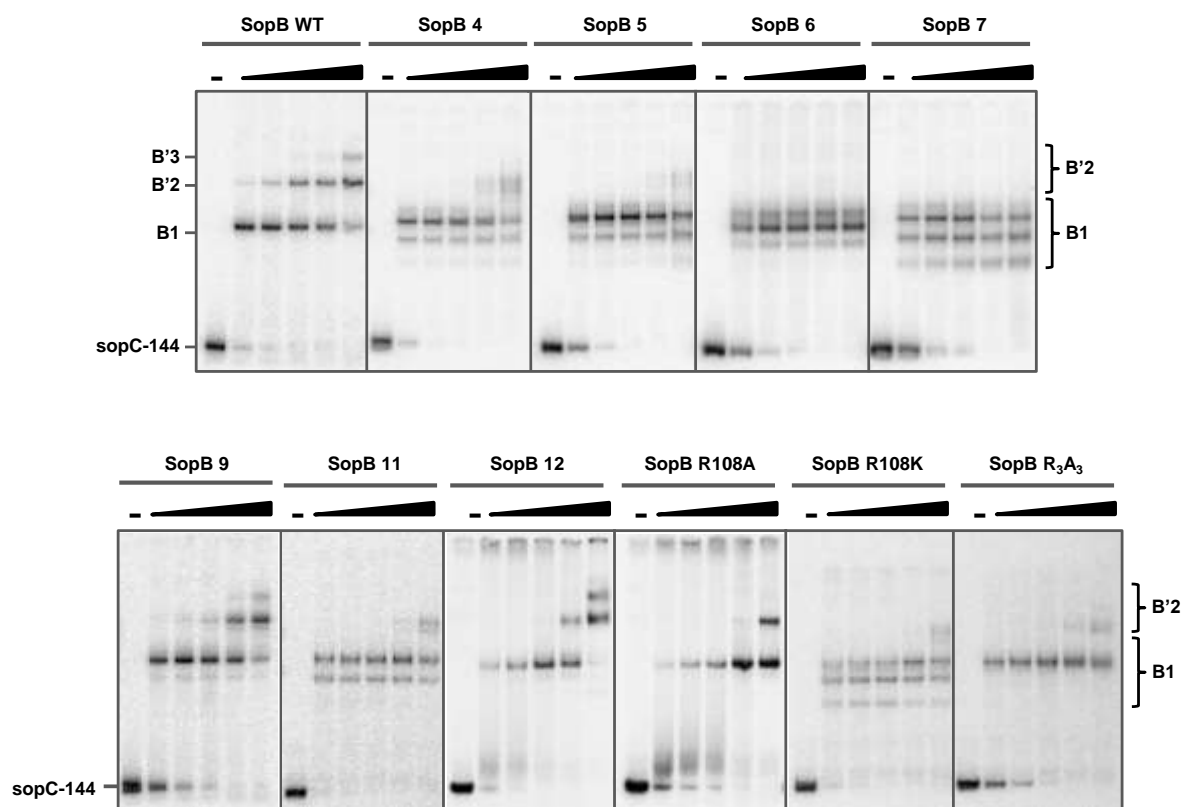


Figure 55 / Mesure par EMSA de la capacité des différents mutants testés à se fixer au centromère et à faire le complexe B2

Des concentrations croissantes de SopB et de ses variants (0, 10, 30, 100, 300, 1000 nM), sont incubées avec 1nM de sonde *sopC* (144 pb) marquée au ³²P. Toutes les réactions contiennent un mélange d'ADN non marqué de sperme de saumon soniqué (100 µg.ml⁻¹) utilisés comme compétiteur d'ADN non spécifique. La position des complexes (B1), (B'2), (B'3) et de la sonde libre est indiquée sur la gauche.

L'interprétation de ces gels à bandes multiples n'est pas directe, on peut arriver à proposer une interprétation si l'on considère qu'il existe des familles de complexe B1 et B'2 comme indiqué sur le côté droite du gel.

Dans le cas des protéines SopB-4, SopB-5, SopB-6, SopB-7 et SopB-R₁₀₈K et selon cette hypothèse, on détecte 3 bandes pouvant correspondre au complexe B1. Ces 3 bandes pourraient illustrer la fixation spécifique de 3 versions de tailles des différents variants de SopB, sur leur centromère. La capacité d'oligomérisation de SopB-4, SopB-5, SopB-R₁₀₈K semble affectée car le complexe B'2 arrive entre 100 et 300 nM alors qu'il est détectable dès 10 nM pour SopB. Les protéines SopB-6 et SopB-7 n'ont plus de complexe B'2, on pourrait en déduire que ces mutants sont défectueux pour leur capacité d'oligomérisation. En recoupant les domaines mutés, on peut envisager que des déterminants d'oligomérisation sont présents entre les résidus 53 et 61 de la protéine SopB. Cependant, dans le cas du mutant SopB-4 dont la délétion comprend celle des mutants SopB-6 et SopB-7, on observe toujours un complexe B'2. Au vu de la complexité des profils de migrations observés, il est donc très difficile d'interpréter et de conclure sur un domaine potentiel d'oligomérisation pour les variants réalisés par délétion de SopB.

Pour la protéine SopB-9, 11 seulement deux bandes pouvant correspondre au complexe B'2, sont détectées. Le mutant SopB-9 n'est pas affecté pour l'oligomérisation car le complexe B'2 arrive dès 10 nM comme pour SopB WT. Par contre, la capacité d'oligomérisation de SopB-11 pourrait être affectée car le complexe B'2 n'arrive qu'entre 100 et 300 nM.

Les mutants SopB 12, SopB-R₁₀₈A et SopB-R₃A₃ ont des profils de migrations sur gel comparables au SopB. Leur capacité d'oligomérisation semble affectée car le complexe B2' n'arrive qu'entre 100 et 300 nM alors qu'il est détectable dès 10 nM de SopB. Les résidus R108 du variant SopB-R₁₀₈A et R₁₂₁-R₁₂₂-R₁₂₃ du variant SopB-R₃A₃, pourraient être des déterminants impliqués dans l'oligomérisation de SopB. De même, les résidus mutés dans le variant SopB-12 comprennent les résidus mutés de SopB-9 et de SopB-11 pour qui on avait respectivement suspecté, pas d'implication ou une implication potentielle. En recoupant les données obtenues pour ces trois mutants, les résidus G₅₂R₅₃-D₅₄ présents dans les variants SopB-8 et SopB-12 puis les résidus W₇₄-S₇₅-G₇₆ présents dans SopB-11 et SopB-12, pourraient également être des déterminants impliqués dans l'oligomérisation de SopB.

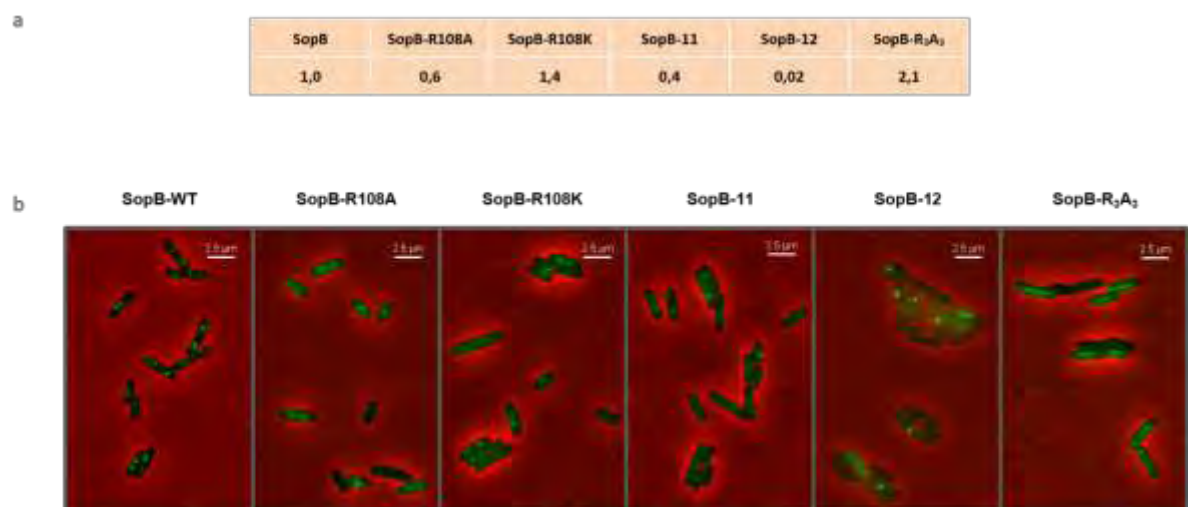


Figure 56 | Les mutants forment des foci moins intenses et délocalisés

- a) Quantification des protéines SopB et des variants de SopB. Des extraits cellulaires de la souche Stellar portant le pDAG114, pJYB234, pJYB236, pJYB237, pAS20, pAS21 et produisant respectivement SopB, SopB-R₁₀₈A, SopB-R₁₀₈K, SopB-11, SopB-12 et SopB-R₃A₃, ont été séparés sur un gel SDS-PAGE. Les gels ont ensuite été soumis à une analyse de Western blot avec un anticorps anti-SopB. Dans chaque piste 0.025 unités de DO₆₀₀ a été déposé. La quantification des protéines a été réalisée avec le logiciel Multi Gauge (Fuji), en prenant comme référence SopB.
- b) Les différentes protéines SopB testées sont indiquées au-dessus de chaque image. Le SopB-mVenus forme des foci intenses avec un positionnement correspondant à celui attendu pour des complexes de partition ayant une ségrégation correcte. Tous les mutants sont capables de former des foci qui sont très faible en intensité comparé au SopB-WT. La grande majorité de la fluorescence diffuse dans tout le cytoplasme.

Cette étude illustre la difficulté d'obtenir des mutants stables et analysables. La région N-terminale de SopB est flexible et la conformation adoptée par cette région semble très importante. En effet, la grande majorité des mutations qu'elles soient par délétion ou ponctuelles, affecte radicalement la stabilité de la protéine. Sur tous les variants testés, nous avons choisis de continuer l'analyse avec les cinq variants, SopB-R₁₀₈A, SopB-R₁₀₈K, SopB-11, SopB-12 et SopB-R₃A₃, dont la capacité d'oligomérisation pourrait être affectée.

C. Localisation des différents variants en microscopie à fluorescence

La protéine SopB, comme les autres protéines ParB, forme des foci discrets et intenses qui colocalisent avec le plasmide portant le centromère. Les protéines se fixant au centromère sont en large excès par rapport au nombre de sites de fixation présents dans la cellule. Dans le cas du plasmide F, SopB est en excès de 15 à 20 fois par rapport au nombre de sites *sopC* (Bouet *et al.*, 2005). Ainsi, ces foci pourraient correspondre à la fixation de SopB sur le centromère puis à la formation d'un complexe de partition "étendu". Dans le cas de mutant ayant perdu leur capacité à s'oligomériser, on s'attend à observer des foci beaucoup moins intenses, voire indétectables, où seules les protéines fixées spécifiquement au centromère seront présentes dans les foci. Tout l'excès de SopB devrait être diffus sur tout le nucléoïde. Pour déterminer si les différents mutants de SopB sont affectés dans leur capacité à former des foci, nous avons construit des dérivés du plasmide mini-F, pJYB234, pJYB236, pJYB237, pAS20, pAS21 et pAS22 avec la protéine fluorescente m-Venus fusionnée en C-terminal avec respectivement, SopB, SopB-R₁₀₈A, SopB-R₁₀₈K, SopB-11, SopB-12 et SopB-R₃A₃. Le plasmide pJYB234 est maintenu au cours des générations dans la souche "Stellar", suggérant la fonctionnalité de la protéine fluorescente. Toutes les protéines fluorescentes sont produites à des concentrations similaires à l'exception du variant SopB-12 qui est très faiblement produit (**Figure 56a**). SopB forme des foci intenses localisés en position médiane ou en positions ¼ et ¾ de la cellule comme cela avait déjà été décrit (**Figure 56b**) (Niki and Hiraga, 1997). Pour les différents variants, on observe des foci avec une intensité beaucoup plus faible et une augmentation du bruit de fond pouvant correspondre à la diffusion des protéines sur le nucléoïde. Les foci des mutants sont délocalisés et certaines cellules ont perdu la fluorescence donc la ségrégation du plasmide est affectée.

L'ensemble de ces SopB variants sont capables de former des foci mais une grosse partie des protéines sont diffuses dans la cellule, suggérant que ces variants pourraient avoir perdu

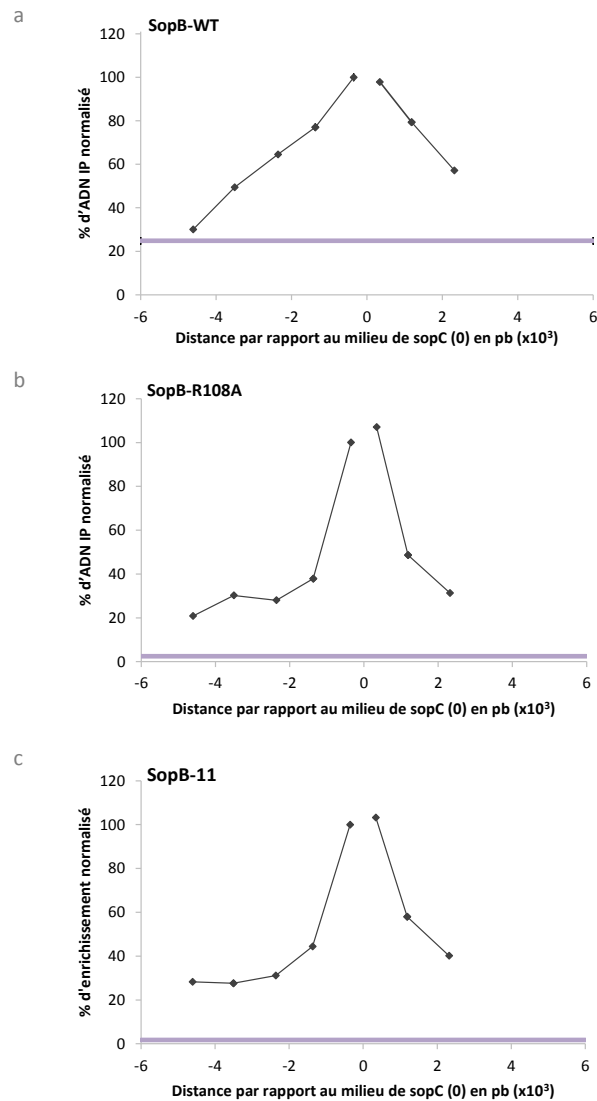


Figure 57 *Étalement de SopB, SopB-R_{108A} et SopB-11 autour du centromère *sopC**

Cette expérience est réalisée sur 3 plasmides, pDAG114, pAS25 et pAS26 produisant les protéines SopB-WT, SopB-R_{108A} ou SopB-11.

Les diagrammes (a), (b) et (c) montrent la quantité d'ADN plasmidique immuno-précipitée en même temps que la protéine SopB ou un de ces variants, en fonction de la distance par rapport au centre du centromère *sopC* (0). La quantité d'ADN plasmidique immuno-précipitée est normalisée par rapport au couple d'oligonucléotide le plus proche de *sopC*, soit AS18. Cet enrichissement est mesuré par PCR quantitative. Le point zéro correspond au centre de la séquence centromérique *sopC*. La droite horizontale mauve montre la quantité d'un locus *terC2*, situé dans la région chromosomique *ter* que l'on peut assimiler à du bruit de fond d'ADN chromosomique immuno-précipité non spécifiquement. Les résultats montrés en (a), (b) et (c) ne représentent qu'une seule expérience, cependant les résultats obtenus en (a) sont similaires à ceux précédemment réalisés dans le **chapitre III des résultats**.

leur capacité de formation du complexe étendu. Ces variants semblent également être affectés pour l'activité de partition, il serait donc intéressant de le vérifier en faisant un test de stabilité. Cependant, chacun des variants a un profil de fluorescence qui lui est propre, ces différences de phénotype pourraient révéler différents niveaux d'atteinte dans la partition.

D. Analyse en qChIP de la capacité des différents variants de SopB à réaliser de l'étalement

En parallèle de l'analyse microscopique, nous avons voulu tester avec la technique de qChIP développée dans le chapitre précédent, si ces cinq variants de SopB étaient affectés pour leur capacité à se fixer en dehors du site centromérique *sopC*. Pour cela, les différents mutants ont été intégrés dans le plasmide mini-F pDAG114 par remplacement allélique, puis transformés dans des cellules DLT1215. Seuls deux mutants ont jusqu'ici pu être obtenus et analysés par cette technique (**Figure 57**). Comme vu précédemment, SopB-WT est capable de se fixer en *cis* sur l'ADN plasmidique à partir du centromère *sopC* et sa capacité de fixation décroît en fonction de l'éloignement de *sopC*. Dans le cas des deux variants testés, on observe des profils similaires entre eux mais très différents de celui obtenu avec SopB-WT. En effet, 60% de l'ADN IP est perdu entre les deux loci flanquant le centromère et les loci suivants. Pour SopB, une perte d'environ 22 % était observée entre ces mêmes loci, ce qui suggère que ces variants pourraient être affectés dans leur capacité à réaliser l'oligomérisation et donc à se fixer en dehors du site centromérique *sopC*.

Dans cette expérience, la quantité d'ADN plasmidique IP est normalisée par rapport au locus défini par le couple d'amorce AS18 (100%). Or, le locus AS18 flanque le site centromérique *sopC* ce qui signifie que la proportion d'ADN mesurée sur ce locus correspond déjà à de la fixation non spécifique. Comme il est impossible d'avoir un couple d'amorce qui amplifie directement le centromère, on ne peut pas avoir de contrôle positif reflétant directement la fixation spécifique de SopB sur son centromère. Grâce aux données de ChIP-seq (**Résultats, Chapitre III, Partie II.C.2**), nous avons montré que chaque position du site centromérique est environ 2.2 fois plus présente que les positions flanquant le site pour le pOX38-B. Dans le cas du pDAG114, les résultats devraient être similaires. Par contre, dans le cas des variants SopB-R₁₀₈A et SopB-11, on ne sait pas qu'elle est la différence entre la proportion de séquences centromériques et les loci qui flanquent directement ces séquences.

Ainsi, la perte de la capacité d'oligomérisation de ces mutants pourrait être largement sous-estimée.

E. Discussion

Cette étude préliminaire visant à caractériser des mutants d'oligomérisation de SopB s'est avérée plus compliquée que prévu et elle s'est heurtée à plusieurs problèmes mais n'en reste pas moins intéressante et prometteuse.

Tout d'abord, nous avons pu voir que la région comprise entre le 52^{ème} et le 118^{ème} aa de SopB est très difficile à manipuler. Les mutations par délétion, dirigées sur quelques acides aminés ou ponctuelles de cette région, ont produit des protéines tronquées lors des purifications, ne permettant qu'une analyse très limitée de leur propriété *in vitro*. Néanmoins, nous avons poursuivi la caractérisation de cinq de ces mutants, SopB R₁₀₈A, SopB-R₁₀₈K, SopB11, SopB12 et SopB-R₃A₃ par des approches *in vivo*. Les analyses en microscopie à fluorescence ont montré que les mutants avaient un profil global similaire. Les variants de ces protéines SopB ne sont plus capables de former des foci discrets, concentrés, localisés aux positions ¼ -¾ de la cellule comme SopB. Ils forment des foci mais la majorité des protéines sont diffuses dans la cellule, suggérant qu'ils seraient affectés pour leur capacité à s'organiser autour du centromère en complexe de partition "étendu".

Chez *B.subtilis*, le système de partition chromosomique ParABS est impliqué dans la sporulation et dans la ségrégation des chromosomes. En effet, l'inactivation de ParB_{B.sub} (SpoOJ) induit un défaut de sporulation et l'apparition de 1% à 2% de cellules anucléés en phase végétatives (Ireton *et al.*, 1994). Un mutant de SpoOJ, SpoOJ93, incapable de réaliser la sporulation et produisant des cellules anucléés, a été isolé (Ireton *et al.* 1994; Piggot and Coote, 1976). Ce mutant résulte du remplacement d'une glycine par une sérine à la position 77 (Lin, 1999), position faisant partie du domaine suspecté d'oligomérisation (Leonard *et al.*, 2004). L'observation en microscopie de la protéine SpoOJ sauvage (SpoOJ-GFP) montre des foci intenses et compacts (Lin *et al.*, 1997; Glaser *et al.*, 1997). Au contraire, le mutant SpoOJ93 ne semble pas pouvoir former de foci et les protéines sont localisées sur le nucléoïde. Ce profil en microscopie a été interprété comme la perte de capacité de ces mutants à s'oligomériser pour créer le complexe de partition "étendu". Contrairement à nos mutants, SpoOJ ne forme plus du tout de foci, deux explications peuvent-être envisagées : (i) soit ce mutant a perdu sa

capacité à se fixer au site *parS* ou (ii) la sensibilité de l'analyse en microscopie ne permet pas de visualiser ces foci de faible intensité. La première explication peut être écartée car ce mutant est toujours capable de se lier au site *parS in vivo* (lin, 1999) et *in vitro* (Breier and Grossman, 2007) comme nos différents variants. La deuxième explication semble la plus probable car en absence d'oligomérisation seul les protéines ParB/SopB fixées sur les sites centromériques sont visualisables. Pour *B. subtilis*, les différents sites *parS* sont espacés de plusieurs Kb ce qui explique qu'ils soient plus difficiles à visualiser contrairement aux 12 sites *sopC* de plasmides F qui sont contigus. Pour ce dernier, la fixation de 10 dimères de SopB sur une courte séquence semble suffire pour la visualisation en microscopie à fluorescence. L'analyse en microscopie des mutants SopB-R₁₀₈A, SopB-R₁₀₈K, SopB-11, SopB-12 et SopB-R₃A₃, devra être complétée par une analyse statistique afin de savoir la proportion des foci délocalisés et celle des foci correctement positionnés aux positions ¼-¾ comme SopB sauvage. Cette étude pourra nous renseigner si la perte d'oligomérisation conduit à un défaut de positionnement des plasmides ou à un défaut ségrégation à proprement dite.

L'expérience de q-ChIP, a pu révéler que les mutants SopB-R₁₀₈A et SopB-11 étaient très affectés pour leur capacité à se fixer sur l'ADN flanquant le centromère *sopC*, ce qui va également dans le sens d'une perte de la capacité à réaliser l'oligomérisation. Cette étude en ChIP devra être complétée par l'analyse des mutants, SopB-R₁₀₈K, SopB-12 et SopB-R₃A₃. Le mutant SpoOJ93 de *B. subtilis* a également fait l'objet d'une étude similaire par la technique de "ChIP on ChIP" (Breier and Grossman, 2007). La comparaison des profils obtenus de SpoOJ sauvage et du mutant SpoOJ93 montre que le pic d'enrichissement n'est pas aussi large avec la protéine variante qu'avec le sauvage, suggérant que ce mutant est affecté dans sa capacité à réaliser l'étalement (Breier and Grossman, 2007). Nos résultats similaires aux leurs, suggèrent fortement que les mutants de SopB testés (SopB-R₁₀₈A et SopB-11), aient eux aussi perdu leur capacité à réaliser le complexe de partition "étendu".

L'importance de l'oligomérisation dans le mécanisme de partition avait été remise en cause par une étude réalisée sur le plasmide P1 (Rodionov and Yarmolinsky, 2004). L'abolition de la propagation des protéines SopB autour du centromère par l'intermédiaire de complexes de type RepA/iterons, n'a pas révélé d'instabilité du plasmide, suggérant que cette propriété n'était pas indispensable à la partition. Au regard des résultats montrés dans le **Chapitre III**, il

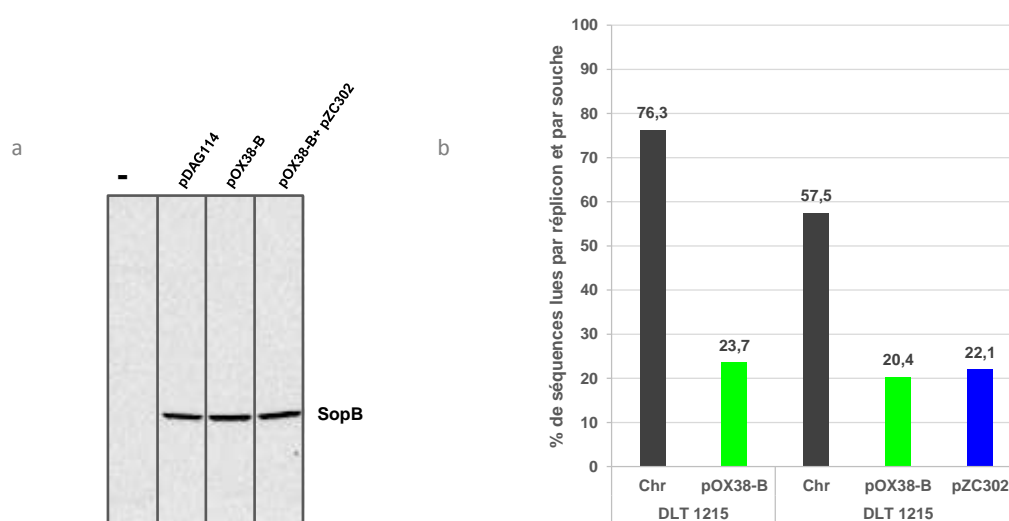


Figure 58 *Répartition de l'étalement de SopB selon les réplicons : chromosomique ou plasmidiques*

- a) Le niveau intracellulaire de protéines SopB produites est similaire avec le pDAG114, le pOX38-B ou le pOX38-B+pZC302. Des extraits cellulaires de la souche DLT1215 seule (-) ou portant le mini-F (pDAG114), le plasmide F-(pOX38-B) ou bien le plasmide F-(pOX38-B) et le pZC302, ont été séparés par SDS-PAGE et soumis à une analyse de Western blot avec un anticorps anti-SopB. Les souches ont été cultivées dans les mêmes conditions de cultures que celle utilisées dans l'expérience de ChIP, dans du LB à 37°C. Dans chaque piste 0.025 de DO₆₀₀ a été déposé.
- b) Le diagramme montre le pourcentage de séquences immuno-précipitées par réplicon et par souche. La souche DLT 1215 contient le plasmide pOX38-B seul ou le plasmide pOX38-B et pZC302. Les deux plasmides contiennent le centromère *sopC* en entier et les souches sont des dérivées de la souche DLT 1215. Cette expérience n'a été réalisée qu'une seule fois.

semblerait que ces barrières physiques de type Rep/itérons ne bloquent pas totalement la propagation de SopB sur le plasmide mais se contente de limiter leur diffusion sur l'ADN. Ainsi, l'expérience réalisée par Rodionov and Yarmolinsky est remise en question. Des tests de stabilité réalisés avec différents mutants d'oligomérisation, pourront donc être envisagés pour savoir si l'oligomérisation des protéines SopB nécessaire à la formation du complexe de partition "étendu", est importante pour la ségrégation des plasmides.

IV. La propagation des protéines SopB sur le chromosome d'*E. coli*

Aucune expérience n'a à ce jour montré un lien direct entre le complexe de partition "étendu" (SopB-plasmide F) et le nucléoïde, ni entre le ségrosome (SopB-SopA-plasmide F) et le nucléoïde.

Afin de savoir si SopB est capable de se propager sur le nucléoïde, j'ai commencé à développer une stratégie en utilisant la technique de ChIP-seq présentée dans le **Chapitre III** des résultats. Grâce à cette technique nous avons été capables de détecter la présence d'un complexe étendu de SopB sur 20 Kb d'ADN plasmidique du plasmide pOX38-B. Lorsqu'on séquence l'échantillon immunoprécipité, on est aussi capable de séquencer de l'ADN chromosomique. Cependant, on est incapable de dire si l'ADN chromosomique immunoprécipité correspond à une immunoprécipitation spécifique résultant de la fixation de SopB sur le chromosome ou bien si cet ADN n'est que le bruit de fond intrinsèque de l'expérience. Pour contourner ceci, nous avons développé une stratégie se basant sur deux faits. Premièrement, SopB est présent en quantité constante dans une cellule car sa production est autorégulée (1000 dimères/cellule) (**Figure 58a**). Deuxièmement, SopB se fixe avec une très forte affinité sur ses sites spécifiques *sopC*, comparé à son affinité pour les séquences d'ADN flanquant. Ainsi, si on insère dans la cellule un plasmide présent en haut nombre de copies et contenant les 12 sites spécifiques *sopC*, on peut espérer déplacer une grande partie des SopB fixés non spécifiquement au niveau du complexe étendu sur les sites spécifiques. Donc, si SopB se fixe uniquement sur le plasmide F (pOX38-B) et pas sur le nucléoïde, on s'attend à une diminution de la proportion de SopB présent sur le pOX38-B au niveau du complexe étendu et ceci au profit du plasmide pZC302 (plasmide haut nombre de copies dérivés du pBR322, ≈20-25 copies/chromosome) lorsque ce dernier est présent. Par contre, la proportion d'ADN chromosomique doit être inchangée.

Dans cette expérience (**Figure 58b**), on observe le cas contraire. Dans une cellule contenant le plasmide pOX38-B et le plasmide pZC302, on observe une légère titration de la quantité de SopB présent sur le plasmide pOX38-B (-3.3%). Par contre, dans ce contexte, on note une diminution de 18,8 % de la quantité d'ADN chromosomique immunoprécipité en même temps que SopB, au profit du plasmide pZC302. En proportion, 60 % des SopB de la cellule sont fixés sur le plasmide contre 40 % sur le chromosome. Nous avons donc montré pour la première fois que SopB est capable de se fixer sur le chromosome *in vivo*. Les 40% de SopB présent sur le chromosome représentent une part non négligeable des SopBs de la cellule, suggérant ainsi un rôle important du nucléoïde. On ne peut pas exclure que la proportion de SopB présent sur le chromosome soit supérieure à 40 % car il est possible que le plasmide pZC302 n'ait pas réalisé la titration complète des SopB présents sur le chromosome. Une étude complémentaire avec un plasmide supérieur en nombre de copies au pZC302 comme un dérivé du plasmide pBSKS (portant *sopC*), pourrait nous permettre de savoir si la proportion des SopB présents sur le chromosome est seulement de 40% ou si cette proportion est supérieure.

L'ensemble de ces données suggère fortement que le complexe "étendu" se propage non seulement en *cis* sur le plasmide mais également en *trans* sur la surface du nucléoïde. Si on met en parallèle ces nouvelles données avec les expériences de microscopie PALM (**Résultat, Chapitre III. Partie II.D.7**), on peut maintenant aller plus loin dans notre modèle : tous les SopB de la cellule sont fortement concentrés autour du centromère en un complexe de partition "étendu" à la fois sur le plasmide et sur le nucléoïde.

V. Perspectives

Ce dernier axe d'étude offre des perspectives intéressantes car il nous a permis de construire des véritables outils pour la compréhension des mécanismes moléculaires impliqués dans le mécanisme global de partition. Lorsque les différents mutants d'interactions avec SopA et d'oligomérisation de SopB seront finis de caractériser, il pourra être envisagé de les utiliser pour essayer de comprendre l'architecture globale du "Ségrosome". Ce dernier se réfère à l'assemblage du complexe de partition "étendu" avec la protéine motrice SopA.

Une collaboration avec le Dr Jun de l'université UCSD à San Diego, a permis de faire des analyses des bactéries avec un dispositif de microscopie couplé à un système de microfluidique (données non publiées de l'équipe). Ce système a permis de visualiser les bactéries, une à une,

dans des micro-chambres. Après induction de la lyse cellulaire, le nucléoïde sort de l'enveloppe membranaire et reste dans la chambre tandis que les constituants membranaires partent dans le système de microfluidique. Dans ces conditions où le flux est constant, le plasmide reste localisé sur le nucléoïde ce qui va dans le sens d'un lien entre le plasmide et le nucléoïde. Dans ce chapitre, nous avons pu mettre en évidence que SopB est capable de se propager sur l'ADN chromosomique, propriété qui pourrait être impliquée dans l'interaction du plasmide avec le nucléoïde.

Il apparaît maintenant primordial de savoir si le lien entre le plasmide et le chromosome nécessite uniquement SopB ou est-ce que ce lien avec le nucléoïde se fait par la protéine SopA. Dans quelle mesure SopA rentre-elle dans l'organisation du complexe de partition "étendu"? A quel moment du mécanisme de ségrégation le complexe de partition a-t-il un rôle ? Est-ce que le complexe de partition "étendu" est capable de réguler la polymérisation de SopA ?

Plusieurs approches pourraient être envisagées pour répondre à ces questions.

Premièrement, en utilisant la technique de microscopie décrite au-dessus et les mutants d'oligomérisation, nous pourrions savoir si le lien qui existe entre le nucléoïde et le plasmide est une conséquence de l'assemblage du complexe de partition "étendu". Avec les mutants d'interaction de SopA, nous pourrions savoir si dans ces conditions le plasmide reste toujours fixé au nucléoïde et donc si SopA participe au lien qui existe entre le plasmide et le complexe de partition.

Deuxièmement, une analyse de sédimentation pourrait également être utile pour continuer à caractériser le complexe de partition "étendu" ainsi que le ségrosome. En effet, une récente étude réalisée sur le plasmide P1, a mis au point un test de sédimentation de sucrose permettant de séparer différentes formes du complexe ParB-ParA-ADN, nommé NAC pour "Nuceoid Adaptator Complex". La formation du NAC requiert de l'ADN, ParB, ParA et de l'ATP (Havey *et al.*, 2012). Le désassemblage de ce complexe requiert la stimulation de l'hydrolyse de l'ATP (Havey *et al.*, 2012; A. G. Vecchiarelli *et al.*, 2013). Nous pourrions grâce à ce test finir de discriminer les mutants capables de réaliser le complexe de partition "étendu" de ceux qui ne le sont plus. Une analyse des changements de topologie de l'ADN du plasmide, avait indirectement montré qu'à forte concentration, SopA était capable de conduire au désassemblage du complexe de partition "étendu" (Lemonnier *et al.*, 2000). Par ailleurs, le

mutant K120Q incapable d'hydrolyser l'ATP n'est plus capable de réaliser ce désassemblage. Le test de sédimentation et l'utilisation des mutants d'interaction de SopB avec SopA, pourraient comme avec l'approche de microscopie décrite ci-dessus, nous renseigner si SopA participe à la formation de ce complexe "étendu". En utilisant le mutant SopA-K120Q et en se plaçant en condition physiologique, nous pourrions également savoir si la perte d'hydrolyse de l'ATP de SopA, comme pour P1 conduit au désassemblage du complexe de partition "étendu".

Avec la technique de ChIP-seq développée dans cette étude et en utilisant ces différents mutants, nous pourrions également arriver aux mêmes conclusions. Cependant, cette technique nous permettrait aussi de préciser et quantifier l'effet de SopA sur la formation du complexe étendu en *cis* et plus précisément si SopA a un rôle dans la proportion de SopB présent sur le plasmide et/ou sur le chromosome ?

Les mutants réalisés dans cette étude pourraient également permettre de poursuivre l'étude initiée dans l'équipe sur l'activation de la polymérisation de SopA par SopB (Bouet et *al.* 2007), en s'intéressant plus particulièrement à la polymérisation en présence du complexe de partition. Notamment, il serait très intéressant de déterminer si l'assemblage de SopB en complexe de partition "étendu", permet une meilleure stimulation de la polymérisation de SopA. Les mutants déficients dans l'oligomérisation pourront être testés, ce qui pourrait confirmer le rôle du complexe de partition "étendu" dans la formation des polymères de SopA. Ces expériences pourront aussi être réalisées avec différents mutants de SopA, déjà caractérisés et déficients pour la fixation à l'ADN ou l'hydrolyse d'ATP (Castaing et *al.* 2008; Libante et *al.* 2001). De même, le test en gradient de sucrose présenté plus haut pourrait également nous renseigner sur le rôle du complexe étendu dans la polymérisation de SopA. Nous pourrions grâce à ce test être capables de détecter l'assemblage des polymères de SopA sur un complexe de partition "étendu" préassemblé.

Une étude ultrastructurale en cryo-électromicroscopie, pourrait être également envisagée afin de visualiser l'architecture du complexe de partition "étendu" et son interaction avec la protéine motrice SopA (Bouet, Ah-Seng, et *al.*, 2007). Dans l'équipe, des polymères de SopA ont pu être imagés en microscopie électronique. Des analyses réalisées en microscopie électronique sur le complexe de partition n'ont révélé que la fixation de SopB sur les sites *sopC* (Hanai et *al.*, 1996). Le complexe de partition "étendu" n'a pu être observé par cette technique,

suggérant que les interactions d'oligomérisation sont perdues au cours de la préparation de l'échantillon. L'avantage de la cryo-électromicroscopie réside dans la congélation ultra-rapide des échantillons, permettant ainsi de conserver une organisation proche de l'état natif des complexes. L'ensemble de ces analyses pourrait se faire avec la protéine SopB pleine taille, mais aussi avec les mutants déficients pour l'oligomérisation, afin de révéler la contribution de l'oligomérisation de SopB dans l'architecture globale du complexe de partition "étendu".

Discussion générale

En utilisant différents modes de fixation à l'ADN et en établissant des interactions protéines-protéines multiples, SopB est capable de se fixer spécifiquement au centromère mais aussi de s'organiser en complexe de partition "étendu" à partir du point de nucléation, *sopC*. L'architecture adoptée par ce complexe de partition n'est connue pour aucun des systèmes de partition ParABS de type I. Les travaux réalisés au cours de cette thèse ont permis de mieux comprendre l'architecture adoptée par ce complexe. En effet, la combinaison d'approche moléculaire et biochimique, a permis de préciser les différentes interactions impliquées dans la formation du complexe de partition, du complexe de partition "étendu" et de comprendre comment il pouvait interagir et réguler les activités de SopA.

La caractérisation des déterminants de spécificité de la séquence centromérique *sopC*, impliqués dans l'interaction avec la protéine ParB, a permis de montrer que toutes les bases de l'inversé-répété sont importantes pour l'interaction. En corrélant ces données avec des études structurales de SopB, un modèle d'interaction entre l'HTH de SopB et le centromère *sopC* a pu être proposé. La formation du complexe de partition, étape indispensable au mécanisme de partition, doit être réalisée exclusivement au niveau du centromère du plasmide, ce qui explique l'importance d'une reconnaissance très spécifique entre SopB et *sopC*. L'importance de l'ensemble des bases de l'inversé-répété donne aussi une explication à la forte interaction ($K_m = 2.5$ nM) observée *in vitro* entre ces deux composants. Une étude réalisée sur la bactérie multichromosomique *B. cenocepacia* a montré que la plupart des bases de l'inversé-répété sont également importantes pour préserver une interaction optimale entre les ParB et les *parS* des différents réplicons, ce qui étend nos résultats de la famille des systèmes de partition Sop aux systèmes chromosomiques (Passot *et al.*, 2012).

La grande spécificité de cette interaction est aussi retrouvée du côté de la protéine SopB. En effet, un nouveau motif, situé en dehors du motif HTH et contenant une arginine, s'est avéré essentiel à l'interaction spécifique avec le centromère, *sopC*. Cette étude a été élargie à tous les ParB des chromosomes principaux et les résultats ont, pour la première fois, révélé une caractéristique conservée dans tout le règne bactérien : les protéines ParB de type Ia contiennent un domaine de liaison au centromère, composé de deux motifs séparés et essentiels. La présence d'un motif de fixation en dehors du motif HTH pourrait également fournir une explication au fait que la protéine chimérique ParB du chr_{III} de *B. cenocepacia* contenant le motif HTH du ParB chr_{II} de *Ralstonia pickettii* 12D, soit toujours capable de fixer

le *parS* de *B. cenocepacia* et vice versa (Passot *et al.*, 2012). En effet, les protéines mutantes conservent une affinité pour leur site *parS*, majoritairement par le corps de la protéine, plus que par leur domaine HTH. Les chromosomes secondaires sont dépourvus du deuxième motif de liaison au centromère caractéristique des chromosomes principaux. Cependant, la présence d'un ou plusieurs résidus en dehors du motif HTH comme pour le plasmide F, pourrait apporter la spécificité de l'interaction entre les ParB et *parS* respectifs. Ainsi, la présence de deux motifs de reconnaissance de *parS* sur ParB et la sensibilité des ParB à des petits changements au niveau des bases de leurs cible *parS* correspondantes, empêche une interaction croisée entre les ParB et *parS* des différents réplicons d'une même bactérie et fournit une explication à l'incompatibilité entre réplicons.

Le cœur de mes travaux de thèse a été de comprendre l'organisation du complexe de partition et du complexe de partition "étendu".

En ce qui concerne l'organisation des complexes de partition de type **I**, les répétitions du site centromérique peuvent s'organiser en répétition directe dans le cas du plasmide F ou non directe dans le plasmide TP228 ou pTAR (**Figure 23**). Les répétitions peuvent être positionnées dans la même orientation (plasmide F) ou dans des orientations différentes (plasmide RK2), mais aussi rassemblées sous forme de cluster de répétition (plasmide pB171). Dans le cas des chromosomes bactériens, les répétitions peuvent être espacées sur toute leur région proximale. Le nombre de répétition et la nature de ces répétitions sont également très hétérogènes (**Tableau 4**). Par conséquent, il n'existe pas de consensus dans l'organisation des centromères de type **I**. En revanche, l'organisation des centromères de type **II** et **III** est semblable entre ces différents types ainsi qu'au sein d'un même type. Cette organisation constituée de répétitions directes et espacées régulièrement, est à l'origine de la structure en superhélice des complexes de partition de type **II** et **III**. Cette superhélice apparaît comme une structure fixe, nécessaire à l'interaction avec le filament de ParM (ParB_{type II}) (**Figure 26 et Figure 27**) (Salje and Löwe, 2008; Aylett and Löwe, 2012).

L'absence de consensus dans l'organisation des centromères de type **I** suggère (i) soit que chaque complexe de partition de type **I** possède une organisation fixe et celle-ci est propre à chacun d'eux (ii) soit qu'il n'existe pas de structure organisée et précise, nécessaire à l'interaction avec la protéine ATPase comme pour les complexes de type **II** et **III**.

Cette première hypothèse me semble peu probable car mis à part leur disparité d'organisation au niveau de leur centromère, le fonctionnement des protéines ATPase de type I est assez semblable, suggérant un mode de fonctionnement commun. De plus, les membres des systèmes de partition de type I qu'ils soient chromosomiques ou plasmidiques, ont en commun la capacité de se propager sur l'ADN plasmidique à partir du centromère, pour former un complexe de partition "étendu", renforçant ainsi la seconde hypothèse.

Le complexe de partition "étendu" avait tout d'abord été décrit comme un étalement des protéines SopB, résultant de leur polymérisation de proche en proche sur l'ADN, à la manière d'un protéo-nucléofilament qui s'initierait au niveau du centromère (Rodionov *et al.*, 1999; Murray *et al.*, 2006). Ce type de propagation suggère une organisation précise, régulière et fixe du complexe de partition "étendu". Cependant, toutes les études ultra-structurales en microscopie électronique ou en AFM (Microscopie de force atomique) qui ont été entreprises, n'ont pas révélé d'assemblage organisé (Hanai *et al.*, 1996; Ringgaard *et al.*, 2007; Pratto *et al.*, 2009). En effet, seules les protéines ParB fixées spécifiquement sur leurs sites de liaison centromériques ont pu être observées et aucune organisation particulière n'a pu être détectée. En conclusion, aucune preuve directe d'un tel protéo-nucléofilament n'a pu être mise en évidence.

La combinaison d'approche moléculaire *in vivo* et *in vitro* avec des expériences de ChIP, de retard de migration sur gel et de microscopie à super-résolution réalisée dans cette étude, a permis de proposer un nouveau modèle dans lequel le complexe de partition "étendu" serait une structure compacte, dynamique et résultant d'interaction stochastique entre SopB et l'ADN avoisinant *sopC* (**Figure 46**).

Au-delà des conclusions apportées dans ce manuscrit, les données recueillies peuvent être utilisées pour faire la modélisation de l'organisation du complexe de partition "étendu". Une collaboration avec l'équipe du docteur Andrea Parmeggiani du département de Physique Théorique du laboratoire Charles Coulomb à Montpellier, est actuellement en cours dans ce sens. Des résultats préliminaires semblent valider le modèle développé dans ce manuscrit.

Les dernières données présentées dans cette étude, suggèrent que SopB est aussi capable de propager le complexe de partition "étendu" sur le nucléoïde tout en restant dans

une zone très concentrée, ce qui pourrait avoir un impact non négligeable dans le mécanisme de partition. Le rôle de cet étalement sur le nucléoïde reste encore à élucider.

Les réactions de partition du plasmide F et P1 ont pu être reconstruites dans un système acellulaire. Dans ce dernier, un tapis d'ADN non spécifique est fixé à la surface d'une lame de microscopie afin de recréer un nucléoïde artificiel (Hwang *et al.*, 2013; Anthony G Vecchiarelli *et al.*, 2013; Vecchiarelli *et al.*, 2014). Par un système de microfluidique, les protéines SopA-ATP et SopB fluorescentes sont injectées sur un tapis d'ADN puis visualisées en microscopie TIRFM. Grâce à cette technique, ils ont observé des comportements dynamiques et des mouvements latéraux pour de nombreux clusters de complexes de partition à la fois pour P1 et F. Ainsi, le nucléoïde servirait de "surface" sur lequel le complexe (SopB-plasmide), pourrait se déplacer. Dans l'expérience de TIRF, lorsque SopB est injectée à concentration équimolaire ou en excès par rapport à SopA, moins de 2% des SopA sont capables de se fixer sur le "nucléoïde" reconstitué. En dépit de la faible densité en SopA, le complexe (plasmide-SopB) est capable de se fixer au "nucléoïde", ce qui indique que SopB pourrait contribuer à réaliser le lien avec le nucléoïde. Notre étude a d'ailleurs suggéré que le complexe de partition "étendu" pourrait s'étendre en *cis* sur le plasmide et en *trans* sur le nucléoïde, fournissant ainsi une explication à ce mécanisme. Par ailleurs, la déplétion progressive des protéines SopA autour du complexe de partition conduit au détachement des clusters de plasmide et à leur mouvement tandis que les protéines SopB restent localisées au niveau du complexe de partition. Ceci suggère fortement que le complexe SopB-plasmide est lié au nucléoïde par l'intermédiaire de SopA-ATP. A ce stade, il apparaît maintenant primordial de savoir si le lien entre le complexe de partition "étendu" (SopB-plasmide F) et le nucléoïde nécessite uniquement SopA ou si ce lien se fait aussi par la protéine SopB ? Dans quelle mesure SopA rentre-elle dans l'organisation du complexe de partition "étendu". Quel est le rôle du complexe étendu et à quel niveau du mécanisme de ségrégation est-il nécessaire ?

Depuis ces dernières années, deux grands modèles de ségrégation se confrontent pour expliquer la séparation des plasmides de type I, le modèle de "réaction-diffusion" et le modèle de "pushing-pulling" (**Introduction, Partie II.A et Figure 14**). Dans le modèle de "réaction-diffusion" initialement présenté par Hunding (Hunding *et al.*, 2003) et les modèles qui en découlent (Fogel and Waldor, 2006; Hester and Lutkenhaus, 2007; Pratto *et al.*, 2008; Ringgaard *et al.*, 2009; Ptacin *et al.*, 2010; Vecchiarelli *et al.*, 2010), un gradient de protéine

ParA serait à l'origine du mouvement des plasmides aux positions $\frac{1}{4}$, $\frac{3}{4}$ de la cellule. Le modèle de "pushing-pulling" quant à lui, repose sur la séparation des plasmides par l'intermédiaire des polymères de la protéine ParA. En effet, les cycles de polymérisation/dépolymérisation de l'ATPase ParA au niveau des complexes de partition, pourraient provoquer le déplacement des plasmides aux pôles opposés de la cellule, soit en les poussant par extension des filaments, soit en les tirant par dépolymérisation des filaments.

Il est très difficile de trancher sur une implication exclusive de l'un ou l'autre des modèles pour expliquer à lui seul la ségrégation des systèmes de type I. En effet, les protéines motrices ParA sont capables de polymériser *in vitro* et d'osciller *in vivo*. Il n'existe pas vraiment de consensus autant dans la capacité de réaliser l'oscillation que dans les conditions de polymérisation pour toutes les ATPase de type I (**Tableau 3**). Pour le plasmide P1, la polymérisation de la protéine ParA n'a pas encore été mise en évidence et toutes les données acquises jusqu'à maintenant semblent indiquer que le modèle "réaction-diffusion" est à l'origine de la ségrégation de ce plasmide. Qu'en est-il pour le plasmide F ? SopA, l'ATPase de ce système, est capable d'osciller dans la cellule. Cette oscillation pourrait générer le gradient de la protéine SopA décrit dans le modèle de "réaction-diffusion". Cependant, SopA est aussi capable de former des polymères *in vitro*, ce qui est aussi compatible avec le modèle de "pushing-pulling".

Il s'agit maintenant de savoir comment l'organisation des complexes de partition s'intègre dans chacun de ces modèles. SopA n'est pas capable de former de filaments en présence d'ADN ce qui pose la question des conditions requises pour former et stabiliser les polymères *in vivo* (Bouet, Ah-Seng, *et al.*, 2007). L'organisation du complexe de partition "étendu" comme structure concentrée et stochastique (**Figure 46**) et les données montrant que SopB est capable de se propager sur l'ADN chromosomique, pourraient expliquer comment SopA polymériserait *in vivo* (**Résultats, Chapitre IV.IV**). En effet, la faculté de concentrer les protéines SopB autour du centromère à la fois en *cis* sur l'ADN plasmidique et en *trans* à la surface du nucléoïde, pourrait permettre d'un part d'organiser l'ADN plasmidique, de le compacter afin de faciliter les étapes de ségrégation ultérieures, mais aussi de créer une zone de forte concentration de SopB, stimulateur de la protéine motrice SopA. Dans le modèle de "réaction-diffusion", la forte concentration des protéines SopB autour du centromère pourrait permettre de stimuler les changements de conformation nécessaires à la protéine SopA pour

réaliser le gradient de concentration impliqué dans le mouvement des plasmides. En effet, l'activité ATPase est essentielle au mécanisme de partition. De plus, la stimulation de l'activité ATPase de SopA est maximale lorsque SopA liée à de l'ADN interagit avec SopB assemblée au niveau du complexe de partition (**Figure 13**) (Ah-Seng *et al.*, 2009). D'après une étude récente, la stimulation de l'activité ATPase pourrait être importante pour la séparation des plasmides frères après réplication. Ainsi, la forte concentration des protéines SopB au niveau du complexe de partition "étendu" pourrait permettre de stimuler l'activité ATPase de SopA et conduire ainsi à la séparation des plasmides frères.

Le modèle de "réaction-diffusion" et le modèle de "pushing-pulling" pourraient aussi coexister *in vivo* mais être impliqués à différents niveaux dans le mécanisme de ségrégation. La polymérisation et l'hydrolyse de l'ATP, pourrait être impliquées dans les étapes précoces de mécanisme permettant la séparation des plasmides frères. On peut imaginer qu'après cette étape précoce de séparation rapide par l'intermédiaire de filaments, les plasmides soient amenés aux positions $\frac{1}{4}$ - $\frac{3}{4}$ de la cellule par un mécanisme de "réaction-diffusion", représenté par l'oscillation de SopA dans la cellule. Par ailleurs, si le mécanisme d'oscillation est le mécanisme permettant la ségrégation des plasmides, on peut également imaginer que ce complexe étendu de SopB sur l'ADN en *cis* et en *trans*, puisse être une cible d'interaction plus large pour le moteur de la partition SopA. Cela augmenterait ainsi la probabilité de reconnaissance du complexe de partition avec le moteur de la partition SopA afin que le plasmide soit bien ségrégué.

Les deux modèles de ségrégation présentés plus haut sont encore incomplets et soulèvent beaucoup de questions. Dans le modèle de "réaction-diffusion", comment se fait la séparation bidirectionnelle des plasmides frères et leur maintien aux positions $\frac{1}{4}$, $\frac{3}{4}$? Dans le modèle de "pushing-pulling", comment les polymères se forment et s'orientent dans la cellule ? Est-ce que la polymérisation de SopA, qui pourrait être initié au niveau de ce complexe étendu, est le mécanisme responsable de la ségrégation globale ou seulement d'une partie ?

Les outils que j'ai développés pendant ma thèse pourront servir à répondre à certaines de ces questions. Grâce à la technique de ChIP-seq et en utilisant des plasmides dépourvus de SopA, nous pourrions savoir si SopA intervient dans la formation du complexe de partition "étendu" et si elle a un rôle dans la proportion de SopB présent sur le plasmide et/ou sur le chromosome ? Lorsqu'ils seront complètement caractérisés, l'utilisation des mutants

d'interaction avec SopA et des mutants d'oligomérisation de SopB, pourraient être des outils précieux dans la compréhension moléculaire du mécanisme de partition.

Matériels et Méthodes

I. Plasmides, Souches, mutagenèse et ADN

Les différents plasmides, souches et oligonucléotides utilisés dans cette étude sont listés dans les (**Tableau 6, 9 et 10**) La séquence de chacun des plasmides construits a été vérifiée par séquençage chez eurofins MWG operon.

A. Mutagenèse par délétion

Les mutants de délétion, SopB Δ 23, SopB Δ 8, SopB Δ 9, SopB-4, SopB-5, SopB-6 et SopB-7 ont été réalisés par PCR. Les oligonucléotides ayant permis de réaliser les délétions sont listés dans le **Tableau 9**. Les fragments PCR contenant les délétions, sont intégrés par clonage via les sites de restriction SpeI et NdeI dans le gène *sopB-inteine*, sur le vecteur d'expression pYAS6 (dérivé de pTYB1). Ce gène est sous le contrôle du promoteur T7. Les vecteurs recombinants, après vérification des séquences des gènes, sont transformés dans des cellules compétentes, C2833 T7 Express (New England Biolabs), pour constituer les souches de production, détaillées en **Tableau 10**. La souche C2833 contient le gène de la T7 polymérase sous le contrôle du promoteur lactose plac, inductible par l'IPTG.

B. Mutagenèse dirigée réalisée par la technique de "quick change"

Les mutations ponctuelles ou de trois acides aminés consécutifs dans SopB ont été introduites par la méthode du kit Stratagene "QuikChange". Cette stratégie est basée sur l'amplification du plasmide entier à partir de deux oligonucléotides complémentaires possédant en leur centre les mutations souhaitées. Cette amplification est réalisée avec l'ADN polymérase PrimeSTAR Max (Takara). Le produit d'amplification est ensuite soumis à l'enzyme DpnI qui clive les brins parentaux méthylés pour enfin être inséré dans une souche DH10B électro-compétente. Les plasmides sont ensuite extraits selon le protocole d'extraction d'ADN plasmidique (kit Plasmid DNA Miniprep de Qiagen) puis introduits dans la souche d'analyse. Les mutants SopB-8, SopB-9, SopB-10, SopB-R₁₀₈A, SopB-R₁₀₈K et SopB-R₃A₃ ont été construits selon ce protocole.

C. Mutagenèse avec le kit "In fusion"

Les mutants SopB-11 et SopB-12 ont été réalisés avec le kit de clonage "In-Fusion PCR Cloning" (Clontech). Des fragments PCR ont été amplifiés à partir des plasmides pAS8, 9 et 10.

Les fragments PCR contiennent entre eux, une homologie de 15 pb à leurs extrémités, ainsi qu'avec le vecteur pYAS6. Les fragments PCR et le plasmide pYAS6 préalablement linéarisé avec les enzymes de restriction NdeI et SpeI sont mis à incuber ensemble selon le protocole prévu par le kit afin d'obtenir après une réaction de recombinaison, un plasmide pleine taille contenant toutes les mutations souhaitées. Le produit de recombinaison est ensuite inséré par transformation dans les cellules "Stellar" contenues dans le kit. Les plasmides sont ensuite extraits puis introduits dans la souche d'analyse.

II. Milieux et antibiotiques utilisés

Toutes les cultures sont effectuées à 30°C ou à 37°C en milieu LB (Luria Bertani : NaCl 10g/L ; extrait de levure 5g/L ; Tryptone 10 g/L ; NaOH 3mM, Thymine 20 mg/L). Les milieux SOB (Tryptone 20 g/L ; Extraits de levure 5 g/L ; NaCl 0.5g/L ; KCl 2.5 mM ; NaOH 1 mM) et SOC (SOB + glucose [2%]_F + MgSO₄ [10 mM]_F) sont utilisés pour la transformation bactérienne par électroporation. Différents milieux M9 (H₂O, Sels M9 1X, CaCl₂ 100 µM ; MgSO₄ 1 mM ; Vitamine B1 ng/ml ; Glycérol 0.4% ou Glucose 0.4% ; Thymine 40 µg/ml ; Leucine 20 µg/ml ou CSA 0.2%, NaOH 8.4 mM) ont été utilisés pour réaliser les courbes de croissance, et le milieu M9-glucose-CSA a été sélectionné pour réaliser l'ensemble des tests de stabilité. Toutes les cultures en milieu solide ont été faites sur milieu LA (milieu LB supplémenté d'agar à 1.5%). Les concentrations d'antibiotiques utilisées sont les suivantes : Gentamycine (Gen) 2.5µg/ml ; Chloramphénicol (Cm) 20 µg/ml (10 µg/ml si la résistance est portée par un plasmide de type miniF) ; Ampicilline (Ap) 100µg/ml ; Kanamycine (Kan) 50 µg/ml ; Spectinomycine (Spc) 30 µg/ml.

III. Transformation

Les plasmides sont introduits dans les souches DH10B, DLT1215 et C2833 par électroporation ou choc thermique. Pour les transformations par électroporation, 5µl de plasmide ont été ajoutés à 45 µl de cellules compétentes. Le mélange a été transféré dans des cuves à électroporation froides et un choc électrique de 2500 V a été réalisé. Les cellules ont été immédiatement reprises dans 450 µl de milieu SOC puis incubées à 37°C pour réaliser l'expression phénotypique. Après 1h30 d'incubation, les cellules ont été étalées sur du milieu LA contenant un antibiotique puis incubées à 37°C sur la nuit.

Pour la transformation au chlorure de rubidium, 1-7 μ l de plasmides ont été ajoutés à 50 μ l de cellules chimiocompétentes. Après incubation 30min dans la glace, les cellules ont été soumises à un choc thermique de 45s à 42°C. L'expression phénotypique est réalisée pendant 1h30 dans 900 μ l de milieu LB à 37°C, puis les cellules ont été étalées sur milieu LA + antibiotique.

IV. Expression et purification de SopA et SopB

A. SopA

La protéine SopA utilisée dans les tests de résonance plasmonique de surface (SPR) et d'activité ATPase, est disponible au laboratoire. Elle a été purifiée sans étiquette (Bouet et *al.* 2007).

Pour les tests de polymérisation en "Dynamic light scattering" (DLS) nous avons utilisé une protéine SopA étiquetée "GSKCK". Cette protéine est purifiée selon un nouveau protocole afin d'améliorer la qualité de la préparation. Elle est produite à partir de la souche DLT2704 (disponible au laboratoire). Les précultures sont réalisées à 37°C dans du milieu LB avec de la Kanamycine (50 μ g/ml) puis diluées 100 fois et incubées à 37°C. A $DO_{600} = 0.3 - 0.4$, la production de SopA est induite avec 0.1% d'arabinose puis l'incubation est poursuivie à 37°C pendant 4h. Les cellules sont ensuite lavées dans du tampon TNE (Tris HCl pH 7.5, 50mM NaCl et 1 mM EDTA), concentrées ($DO_{600} \approx 200$) et resuspendues dans du tampon A (20 mM Tris HCl pH 7.8, 100 mM NaCl, 1 mM EDTA, 20% sucrose). Après ajout de lysozyme (200 μ g.ml⁻¹), les cellules sont immédiatement congelées dans de l'azote liquide. Après décongélation pendant une nuit dans la glace, les cellules sont centrifugées 90 min à 4°C, à 100 000 g. Le culot, contenant SopA, est resuspendu dans le tampon A, soniqué puis centrifugé à 40 000 g pendant 25 min à 4°C. La purification est ensuite réalisée selon le protocole décrit dans Bouet et *al.* 2007. Toutes les étapes de purification sont réalisées à 4°C et contrôlées sur gel SDS-PAGE. Le surnageant (Fraction I S) contenant SopA est précipité au sulfate d'ammonium (0.32 mg.ml⁻¹) pendant 30 min puis centrifugé pendant 20 min à 12 000g. Le culot est resuspendu dans 2,5 ml de tampon S (40 mM Hepes-KOH pH 7.8, 75 mM KCl, 0.1 mM EDTA 10% glycérol) et dialysé sur colonne de dessalage (PD10) avec du tampon S. L'élution de la dialyse est chargée sur une colonne de DEAE Sephadex (30 ml, Pharmacia) et éluée avec un gradient de 75-600 mM KCl dans du tampon S. Les fractions contenant SopA sont rassemblées, diluées et constituent la

fraction **II**. La fraction **II** est ensuite chargée sur colonne d'Héparine et éluée avec un gradient de 100 mM – 800 mM de KCl dans du tampon S. Les fractions contenant SopA sont rassemblées (Fraction **III**) puis précipitées au sulfate d'ammonium (0.32 mg.ml⁻¹) pendant 30 min puis centrifugées pendant 20min à 12 000 g. Le culot est resuspendu dans 2 ml de tampon de tampon GF (40 mM Hepes-KOH pH 7.8, 400 mM KCl) puis injecté sur une colonne de chromatographie d'exclusion (Superdex 200, Pharmacia) préalablement équilibrée avec le tampon GF. Les fractions contenant SopA sont rassemblées (Fraction IV) puis chargées sur une colonne échangeuse d'anions de type mono-Q (GE Healthcare) et éluées avec 250-750 mM de KCl dans du tampon S. Les différentes fractions V collectées sont dosées au nanodrop, puis congelées avec de l'azote liquide et stockées à -80°C.

B. Purification de SopB étiquetée intéine

SopB-G₃₂₄ et les variants de SopB étiquetés intéines, sont purifiés à partir de la souche de production C2833 suivant le protocole adapté de (Ah-Seng et *al.* 2009). Les souches sont cultivées toute la nuit à 30°C dans un milieu LB avec de l'ampicilline (100 µg.ml). Les précultures sont diluées 100 fois et incubées à 42°C. A DO₆₀₀ = 0.5-0.8, la production de SopB-intéine est induite avec 0.4 mM d'isopropyl β-D-1-thiogalactopyranoside (IPTG) puis les cellules sont incubées à 30°C, pendant 4h, pour SopB-G₃₂₄ et SopB-G₃₂₄-Δ₂₃-inteine, et à 20°C pendant 16 à 20h pour toutes les autres variants de SopB. Les cellules sont ensuite lysées par sonication (15 min, 50% de temps actifs, puissance de sortie 4) dans du tampon B (20 mM Tris HCl pH 7.8 ; 500 mM NaCl ; 1 mM EDTA) contenant 10 µg.ml⁻¹ de lysozyme. Le surnageant (Frl), obtenu par centrifugation à 4°C pendant 20 min à 20 000g, est chargé sur une colonne de chitine (5 ml), préalablement équilibrée dans du tampon B. Puis la colonne est lavée successivement avec 15 ml de tampon B, 15 ml de tampon B contenant 1M de NaCl et 15 ml de tampon B. La colonne est ensuite incubée avec 15 ml de tampon B contenant 50 mM de DTT et stockée à 4°C pendant 16 à 20h pour permettre l'obtention d'une protéine non étiquetée après le clivage de l'étiquette intéine. Les protéines sont éluées de la colonne de chitine avec 15 ml de tampon B puis concentrées dans 2 ml pour constituer la fraction **II**. La concentration des protéines est réalisée soit (i) par rétention sur filtre, (Amicon 15-30 K, Millipore) soit par (ii) précipitation au sulfate d'ammonium. Pour la précipitation, les fractions sont incubées à 4°C, sous agitation douce, avec 0.40 g/ml de sulfate d'ammonium pendant 30-60 min. Elles sont ensuite centrifugées 20 min à 20 000 g et le culot est congelé dans l'azote liquide. Après décongélation

du culot, les protéines sont resuspendues dans 2 ml d'un tampon contenant 40 mM Hepes-KOH pH 7.8, 400 mM KCl.

La fraction **II** concentrée (2 ml) est centrifugée à (14 000 rpm) pendant 1 min pour se débarrasser des agrégats puis injectée sur une colonne de chromatographie d'exclusion (Superdex 200, Pharmacia) préalablement équilibrée avec 40 mM Hepes-KOH pH 7.8 et 400 mM KCl. Les fractions contenant la protéine SopB sont regroupées (fraction **III**) puis diluées pour obtenir une concentration de 100 mM en KCl afin de pouvoir directement charger la fraction **III** sur une colonne échangeuse de cation de type mono-S (GE Healthcare). La protéine SopB est ensuite éluée grâce à un gradient de 100 à 600 mM KCl dans du tampon S. Dans la majorité des cas, la protéine SopB est éluée à 350 mM de KCl. Les différentes fractions IV (FrIV) collectées sont dosées au nanodrop ou par la méthode de "Bradford protein assay" (BIORAD), puis congelées avec de l'azote liquide et stockées à -80°C. Les 3 étapes de purification permettent d'atteindre entre 90 et 95% de pureté.

Afin de contrôler la production et la purification des protéines SopB, des aliquots ont été prélevés avant l'induction à l'IPTG (N), en fin de culture (I), dans le culot de la 2^{ème} centrifugation (C), dans son surnageant (S), après élution de la colonne de chitine (Fr**II**), de la colonne de chromatographie d'exclusion (FrIV) et de la colonne mono-S (FrV). Puis, tous les aliquots sont centrifugés puis re-suspendus dans du tampon de charge (10% glycérol ; 5% β -mercaptoéthanol (BME) ; 2.3% SDS ; 62.5 mM Tris-HCL pH 6.8 ; 0.25% bleu de bromophénol) ou dilués dans ce tampon, afin d'obtenir un équivalent cellulaire de DO600=0,03 par piste. Les échantillons migrent pendant 50 min à 200 V sur des gels SDS-PAGE 4-12% d'acrylamide (NuPage- Invitrogen), avant d'être révélés par une solution de bleu de coomassie (Instant blue-Expedeon) après 1h d'incubation à température ambiante.

V. Test de retard de migration sur gel

A. Préparation des sondes

Les sondes radiomarquées utilisées pour les expériences d'EMSA sont préparées soit par amplification PCR en présence de [α -³²P] dATP, soit par hybridation de paires d'oligonucléotides et réaction de kinase en 5' avec du [γ -³²P] ATP. La sonde *repE114* (114pb) est amplifiée à partir du plasmide, pDAG114, avec les oligonucléotides *repE1* et *repE2*. La sonde

sopC144 (144 pb) est amplifiée à partir du plasmide, pJYB55, avec les oligonucléotides *sopC22* et *sopC23*. Les sondes sont ensuite purifiées avec un kit de purification d'ADN (Promega).

Les autres sondes *sopC*, listées dans le **Tableau 7**, ont été obtenues par hybridation de paires d'oligonucléotides synthétiques complémentaires. Le mélange a été chauffé à 95°C pendant 15min puis refroidi lentement afin de permettre une hybridation correcte des oligonucléotides complémentaires lors de la descente en température. Les sondes obtenues ont ensuite été marquées radioactivement par phosphorylation des extrémités 5'. Le mélange réactionnel de 50 µl a été le suivant : 6.25 µl de sonde froide à 0.4 µM ; 5 µl d'ATPγ³²P à 10 mCi/ml ; 2.5 µl de tampon PNK 10X; 1 µl de PNK à 10 U/µl (NEBiolabs); H₂O qsp 25 µl. La réaction est incubée 45 min à 37°C, puis inactivée 10 min à 68°C. Les sondes marquées sont ensuite purifiées sur colonne G25 (GE Healthcare).

A. Réactions d'interaction *in vitro*

Les réactions (10 µl) contiennent 0.3 à 1 nM d'ADN marqué au ³²P dans un tampon composé de 30 mM Hepes-KOH (pH 7,5), 100 mM de KCL, 10% de glycérol, 1 mM de dithiothreitol (DTT), 100 µg.ml⁻¹ de BSA (sérum albumine bovine) et 100 µg.ml⁻¹ d'ADN de sperme de saumon soniqué. Les mélanges réactionnels avec les concentrations en protéine voulues, sont assemblés dans la glace puis incubés pendant 15 min à 30°C. Ils sont ensuite analysés par électrophorèse dans un gel d'acrylamide (6%) TGE (50 mM Tris glycine pH = 7.8, 1 mM EDTA), pré-migré à 75 volts (V) pendant 30 min à 4°C. Les échantillons sont chargés à 75V puis la migration est effectuée durant 2 à 3h50, à 180 Volt et à 4°C. Après fixation dans une solution contenant 10% d'éthanol et 10% d'acide acétique puis séchage sous vide sur papier Whatman, les gels sont exposés sur écran PhosphorImager (Fuji FLA3000). L'analyse des résultats se fait avec le logiciel Multi Gauge (Fuji).

VI. Western-Blot

A DO₆₀₀ = 0.5, les cultures bactériennes sont centrifugées pendant 5 min à 14000g et resuspendu dans 150 µl du tampon de charge (10% glycérol ; 5% β-mercaptoéthanol (BME) ; 2.3% SDS ; 62.5 mM Tris-HCL pH 6.8 ; 0.25% bleu de bromophénol). Les échantillons sont soniqués pendant une demi-seconde, l'étape est répétée 3 ou 4 fois. Les extraits protéiques sont incubés 5 min à 95°C, vortexés 30 secondes, et déposés sur gel SDS-PAGE 4-12%

d'acrylamide (NuPage, Invitrogen) pour une migration de 50 min à 200 Volt. Le gel est ensuite transféré sur membrane de Nitrocellulose avec l'iBlot d'Invitrogen pendant 7.5 min à 20 volts. La membrane est saturée avec 50 ml de tampon TBS 1X-5% lait (Tris-HCl pH 4.4 20mM ; NaCl 137 mM ; 0.05% Tween et 5% de lait) pendant 2h. La solution est remplacée par 50 ml de tampon TBS 1X supplémenté avec du sérum anti SopB et mis à incuber toute la nuit à 4°C sous agitation douce. La membrane est ensuite lavée avec la solution de TBS 1X- lait 0.5% pendant 20 min et l'étape est répétée 2 fois. L'anticorps secondaire IgG anti-rabbit/HRP, dilué au 1/2000 dans 20 ml de tampon TBS-lait 0.5%, est mis à incuber 1h30 avec la membrane. Cette dernière est de nouveau lavée 3 fois, avec 50 ml de tampon TBS 1X - 0.5% lait, pendant 20 min. Puis un dernier lavage est effectué avec 50 ml de tampon TBS 1X - Tween 0.1% sans lait. Pour finir, la membrane est incubée avec la solution d' "ECL prime" (Enhanced ChemiLuminescent, GE Healthcare) pendant 5 min, puis révélée avec le Biolumager LAS-4000 (Fujifilm).

VII. Footprint *in vitro*

Pour réaliser les expériences de footprint à la DNase I, 10 nM de sonde ADN radiomarquée ont été utilisés. Les sondes ont un seul site de liaison de SopB de 16pb (*sopC223*) ou deux sites (*sopC-266*). Les réactions (20 µl) contiennent 10 nM de sonde radiomarquée dans un tampon composé de 30mM Hepes-KOH (pH 7,5), 100mM de KCL, 10% de glycérol, 1mM de dithiothreitol (DTT), 100µg.ml⁻¹ de BSA (sérum albumine bovine) et 100µg.ml⁻¹ d'ADN de sperme de saumon soniqué. Les mélanges réactionnels sont assemblés dans la glace puis incubés pendant 15 min à 30°C. 2 µl d'une solution de DNase I à 1 U.ml⁻¹ (BioLabs) ont ensuite été rajoutés aux réactions qui sont mises à incuber 1 min à 37°C. Les réactions sont arrêtées par ajout de 5 µl d'une solution STOP (95% formamide 95%, 10 mM Na₃EDTA, 0.02% de bleu de bromphenol, 10 mM NaOH), suivis d'une précipitation à l'éthanol. L'ADN est resuspendu dans 10 µl d'H₂O puis incubé 5 min à 95°C. 6 µl de chaque réaction sont déposés sur un gel de polyacrylamide-urée 6%. Les gels sont séchés, puis exposés sur des écrans photo-stimulables (Fuji). Les profils de coupure à la DNaseI sont ensuite révélés par lecture des écrans sur le typhoon Trio (Fuji-GE Healthcare).

VIII. Test d'activité ATPase

L'activité ATPase de SopA est mesurée par la quantification d'ADP formé lors de l'incubation de 2 μM de SopA dans 10 μL de tampon SCAB (50mM Hepes-KOH pH 7.8, 50 mM MgCl_2 , 1 mM DTT, 10% glycérol) contenant 1 mM $[\alpha\text{-}^{32}\text{P}]$ ATP. Dans la réaction sera ajouté comme indiqué, de l'ADN non spécifique, (plasmide pBSKS) ou de l'ADN spécifique, (plasmide pJYB57 contenant *sopC*) à 100 $\mu\text{g}.\text{ml}^{-1}$. La protéine SopB est ajoutée à une concentration de 2 μM . Les mélanges réactionnels sont incubés à 37°C pendant 4H et la réaction est ensuite arrêtée par l'ajout de 10 μL d'un tampon (3 mM ATP ; 3 mM ADP ; 50 mM EDTA). 1.5 μL de chaque essai est déposé sur une plaque TLC et la chromatographie se fait par capillarité dans du tampon composé de 1M d'acide formique et 0.5 M de LiCl. Les plaques sont séchées et analysées au phosphorImager (Fuji FLA-3000). La quantification de la radioactivité est ensuite réalisée par traitement des images avec le logiciel Multi Gauge (Fuji).

IX. Test de stabilité

Les tests de stabilité ont été réalisés comme décrit dans (Dubarry et *al.* 2006). Toutes les expériences se font à partir de colonies de cellules "fraîchement" transformées avec le plasmide à tester. Les différents allèles de SopB, exprimés dans des plasmides dérivés de pSC101 sous le contrôle du promoteur *plac* sont introduits dans des souches MC1061 contenant déjà le plasmide mini-F à tester. Ce plasmide est un dérivé du pDAG114 délété de *sopA* et d'une grande partie de *sopB*, ne permettant pas la production de la protéine fonctionnelle. Des précultures des souches ont été réalisées en milieu M9-glucose-CSA à 30°C en présence des antibiotiques nécessaires au maintien des deux plasmides. Ces précultures ont été diluées 400 fois le matin dans le même milieu et incubées à 30°C jusqu'à une $\text{DO}_{600}=0.25$.

Les cultures sont diluées 52 000 fois (19 générations) dans du milieu M9-glucose-CSA, sans antibiotique (donc sans pression de sélection pour le pDAG209), et mises à 30°C jusqu'à ce qu'elles atteignent la même densité cellulaire. La dilution et la croissance sont répétées durant 30 générations. Pour déterminer la proportion de cellule possédant le plasmide pDAG209, des échantillons des cultures sont prélevés au début de l'expérience et à chaque dilution des cultures. Ces échantillons sont étalés sur du milieu LA sans antibiotiques et mis à 37°C toute la nuit. Les colonies obtenues ont été répliquées sur velours puis mises sur milieu

sélectif (LA+ antibiotique correspondant à la résistance conférée par le pDAG209). Le pourcentage de plasmides perdus par génération (L) est calculé suivant l'équation suivante : $L = 1 - (F_f/F_i)^{1/n} \times 100$, où F_i est la fraction de cellules portant initialement le plasmide et F_f est la fraction des cellules portant le plasmide après n générations en milieu non sélectif.

X. Microscopie à fluorescence (Chapitre IV)

Les cellules Stellar sont transformées avec les plasmides dérivés du mini-F, pJYB234, pJYB 236, pJYB237, pAS20 et pAS21 contenant la protéine fluorescente m-Venus fusionnée en C-terminal avec respectivement, SopB-WT, SopB-R₁₀₈A, SopB-R₁₀₈K, SopB-11 et SopB-12. Les transformants sont mis en culture une nuit à 30°C dans un milieu M9-glucose-CSA. Les cultures sont diluées au 400^{ème} dans le même milieu puis incubées à 30°C jusqu'à l'obtention d'une DO₆₀₀ de 0.3. Les cultures sont ensuite déposées sur une lame recouverte d'un coussin d'agarose à 1% préparé avec du milieu M9-1X. Les cellules sont examinées en contraste de phase (objectif x100) par microscopie à épi-fluorescence à l'aide du microscope Nikon Ti –E. Les lames sont maintenues à 30°C dans une chambre thermostatée durant toute la durée de l'expérience. L'acquisition des images est réalisée avec la caméra OrcaR2 (Hamamatsu) (expo 1 s, contraste de gain 144), le traitement des images est effectué avec le logiciel NIS-élément.

XI. Résonance plasmonique de surface

Les études de liaison et de cinétique sont réalisées avec le Biacore 3000. Les étapes d'immobilisation des sondes ADN biotinylées sur un "sensor chip" se font jusqu'à l'obtention d'une immobilisation (de l'ordre de 500 RU) équivalente des sondes ADN (136 pb), non spécifiques (référence) et spécifiques (contenant un motif *sopC*). Les analyses d'interaction sont effectuées avec de multiples injections de différentes concentrations de protéines à 25°C. Les échantillons, dilués dans du tampon BD (20 mM Hepes-KOH pH 7.4 ; 100 mM KCl ; 10 mM MgCl₂ ; 50 µg.ml⁻¹ d'ADN compétiteur de sperme de saumon soniqué, 1 mM de DTT), sont injectés à 10 µL.min⁻¹ sur la surface du "sensor chip" pendant 180 s en présence d'1 mM d'ATP. Les données obtenues avec la sonde de référence (ADN non spécifique) ont été soustraites de celles obtenues avec la sonde *sopC* (ADN spécifique) pour permettre de visualiser directement et uniquement les interactions spécifiques des protéines à la sonde *sopC*. Ces expériences sont

faites en collaboration avec Frédéric Lopez de la plateforme protéomique de l'Institut Louis Bugnard (CHU Rangueil, Toulouse).

XII. Expérience de diffusion de lumière dynamique (DLS)

Les protéines SopA (étiquetée "GSKCK") et SopB sont préalablement centrifugées 10 min à 20 000g à 4°C afin de se débarrasser de tous les agrégats. SopA est ensuite placée dans du tampon PA filtré, (50 mM Hepes-KOH pH 7.5 ; 50 mM KCl ; 2 mM MgCl₂ ; 1 mM DTT) en présence ou en absence de SopB, puis la réaction est centrifugée 3 min à 4°C. L'ATP (1 nM) est ajouté au surnageant et le mélange réactionnel est transféré dans une cuvette de quartz. Celle-ci est placée dans un photomètre Dynapro System et exposée à un laser émettant à 824 nm. La déviation de la lumière à 90° est mesurée toutes les 10s et analysée par le programme Dynamics V6. L'intensité mesurée est proportionnelle à la taille et à la concentration des particules déviant la lumière. Les expériences sont réalisées à 20°C.

XIII. Immunoprécipitation de Chromatine (ChIP)

A. Purification des anticorps anti-SopB

100 µg de protéine SopB purifiée sont incubée 5 min à 95°C avec un tampon de charge 4X (NuPage) puis déposées sur gel SDS-PAGE 4-15% d'acrylamide (mini-PROTEAN TGX-Biorad), mono-puits. Après migration pendant 30 min à 200 V, le gel est transféré sur membrane de Nitrocellulose avec l'iBlot d'Invitrogen pendant 7.5 min à 20 volts. La membrane est d'abord rincée 10 min avec 20 ml de tampon (tris-HCl pH 4.4, 20 mM ; NaCl, 137 mM) puis bloquée avec 20 ml de tampon TBS 1X-10% lait (tris-HCl pH 4.4, 20 mM ; NaCl, 137 mM ; 0.1% Tween et 10% de lait) pendant 1h à température ambiante. La solution est remplacée par 6 ml de tampon TBS 1X (tris-HCl pH 4.4, 20 mM ; NaCl, 137 mM ; 0.1% Tween) supplémenté avec 500 µl de sérum anti SopB et mis à incuber toute la nuit à 4°C sous agitation douce. La membrane est ensuite lavée 5 fois avec le tampon TBS 1X (tris-HCl pH 4.4, 20mM ; NaCl, 137 mM ; 0.1% Tween) pendant 10 min puis 2 fois avec du tampon Na-phosphate (0.1 M Na₂PO₄, 0.1 M Na₂H₂PO₄) pendant 5 min. La membrane est ensuite éluée avec 400 µl de glycine 0.1M pH 2.2. L'éluât est ensuite mélangé avec 40 µl de Na₂HPO₄, 1 M et 0.02% d'azide de sodium avant d'être dosé au NanoDrop (Thermo Scientific).

B. Culture et traitement au formaldéhyde

Les expériences d'immunoprécipitation de chromatine sont effectuées sur des souches DLT1215 contenant le plasmide mini-F (pDAG114) ou le pOX38-B. Des précultures des différentes souches sont réalisées en milieu LB à 37°C en présence des antibiotiques nécessaires au maintien des plasmides. Ces précultures sont diluées 200 fois le matin dans le même milieu et incubées à 37°C. A $DO_{600}=0.6$, 1% de formaldéhyde est ajouté à 80ml de cette culture puis le tout est mis à incuber à température ambiante pendant 30 minutes, sous agitation (180 rpm). La réaction est arrêtée en ajoutant de la glycine (500 mM final) et en incubant 15 min à température ambiante. Les cellules sont ensuite centrifugées 10 min à 6 000 rpm. Le culot est d'abord rincé avec du tampon Tris salin (20 mM de Tris-HCl, pH 7.6 ; 100 mM de NaCl), resuspendu dans 1ml de tampon de lyse (10 mM de Tris-HCl pH 7.8 ; 10 mM d'EDTA ; 100 mM de NaCl ; 20% de sucrose ; 10 mM d'EDTA) contenant 1 mg/ml de lysosyme et mis à incuber à 30°C. Après 30 minutes d'incubation, 1ml de tampon IP (50 mM d'Hepes-KOH, pH 7.5 ; 150 mM de NaCl ; 1 mM d'EDTA ; 1% de triton X-100 ; 0.1% de sodium deoxycholate ; 0.1% de SDS ; 1 mM phenylmethylsulfonyl fluoride (PMSF)) est rajouté puis séparé en deux fractions de 1 ml dans des tubes de 1.5 ml.

1. Sonication

Les fractions sont ensuite soniquées dans un Biorupteur (Diagenode), à haute puissance, durant 15-18 cycles de 30 secondes séparés par 30 s de repos, jusqu'à l'obtention de fragments d'ADN compris entre 500 et 1 000 pb. Les tubes sont ensuite centrifugés 30 min à 20 000 g et à 4°C, pour se débarrasser des débris cellulaires. Le surnageant des deux tubes est récupéré, rassemblé (2 ml) et représente l'échantillon de départ, appelé input. 500 µl sont prélevés pour constituer la fraction Non ImmunoPrécipitée (NIP). La sonication est vérifiée sur cet échantillon. Pour cela, celui-ci est digéré 1H à 37°C à la RNase A (1 µg/µL) puis 2H à 37°C avec de la protéinase K (0.5 µg/µL) après ajout de 1% de SDS et incubé une nuit à 65°C pour reverser le cross-link. L'ADN est alors purifié par extraction au chloroforme/alcool isoamylique (24 :1, pH 8), suivi d'une précipitation par ajout de 0.7 volume d'isopropanol froid et 40 µg/ml de glycogène (Roche). L'échantillon est incubé toute la nuit à -20°C. Après 30 min de centrifugation à 16 000 g et à 4°C, le culot est rincé à l'éthanol 70%, puis séché avant d'être repris dans 100 µL d'eau milli-Q. L'ADN est ensuite analysé sur gel d'agarose 1% (TBE 0.5X), pour vérifier la taille des ADNs, entre 500 et 1 000 pb.

2. Immunoprécipitation et purification de l'ADN

Toutes les étapes d'immunoprécipitation et de blocage se font à 4°C sur roue et ne sont réalisées qu'avec des billes magnétiques recouvertes de protéine A (Ademtech), préalablement bloquées. Le blocage se fait par incubation des billes pendant une nuit dans le tampon IP supplémenté en BSA (0.1 µg/µl) et tRNA (1 µg/µl). Entre chaque étape, les billes sont immunoprécipitées avec le kit modulo (Ademtech).

1 ml d'input est pré-incubé avec 50 µl des billes magnétiques préalablement bloquées pendant 1h. Cette pré-incubation permet d'enlever les complexes ADN-protéines ayant une affinité non-spécifique pour les billes et diminuer ainsi le bruit de fond intrinsèque de l'immunoprécipitation ultérieure. Le surnageant de cette pré-incubation est prélevé et divisé en deux : 500 µl constituent la fraction IP qui sera immunoprécipitée avec les anticorps anti-SopB et les 500 µl restant seront traités en parallèle et serviront de contrôle négatif de cette immunoprécipitation, appelé IP(-). Les deux fractions sont incubées une nuit, en présence ou en absence de 7 µg d'anticorps anti-SopB. Les deux fractions sont ensuite incubées pendant 2h avec 25 µl de nouvelles billes magnétiques fraîchement bloquées. Les billes sont ensuite précipitées et lavées 10 min avec 750 µl de différents tampons : 1 fois avec le tampon IP, 2 fois avec le tampon **II** (50mM Hepes-KOH, pH 7.5 ; 500 mM NaCl ; 1 mM EDTA ; 1% triton X-100 ; 0.1% sodium deoxycholate ; 0.1% SDS ; 1 mM PMSF), 1 fois avec le tampon **III** (10 mM Tris-HCl pH 7.8 ; 250 mM LiCl ; 1 mM EDTA ; 0.5% Nonidet-P40 ; 0.1% sodium deoxycholate) et 1 fois avec du tampon TE (10 mM Tris pH 7.5, 1 mM EDTA). Les billes sont ensuite éluées 1h à 37°C avec 300 µl de tampon d'éluion (50 mM Tris-HCl pH7.5 ; 10 mM EDTA ; 1% SDS) contenant 1 µg/µl de RNase A suivis de 2h d'incubation avec 0.3 µg/µl de protéinase K puis immunoprécipitées. Le surnageant est incubé une nuit à 65°C afin de reverser le cross-link. L'ADN est alors purifié par extraction au phénol/chloroforme/alcool isoamylique suivi d'une précipitation par ajout de 0.7 volume d'isopropanol froid et 40 µg/ml de glycogène (Roche). L'échantillon est incubé toute la nuit à -20°C. Après 30 min de centrifugation à 16 000 × g et à 4°C, le culot est rincé à l'éthanol 70%, puis séché avant d'être repris dans 100 µL d'eau milli-Q.

L'ADN est ensuite resuspendu dans de l'eau milliQ pour être quantifié par qPCR (séquences des amorces dans le **Tableau 8**) ou dans 10 mM de Tris-HCl, pH 8.5 pour l'analyse en séquençage haut débit.

XIV. PCR quantitative

A. Choix des différentes amorces

Les différentes matrices oligonucléotidiques de 20 pb sont choisies en utilisant le logiciel gratuit d'invitrogen (<http://tools.invitrogen.com/content.cfm?pageID=9716>) et listées dans le **Tableau 8**. Elles ont une température d'hybridation d'environ 60°C, un pourcentage GC compris entre 50 et 60. Chaque couple d'amorce réalise des amplifications de petites tailles (moins de 200 pb). Un Blast contre *E. coli* est réalisé avec le logiciel primer blast sur le site internet du NCBI, pour vérifier que les amorces sont spécifiques des séquences voulues. Les couples d'amorces sont localisés de part et d'autre du centromère puis répartis sur l'ensemble des plasmides étudiés. Chaque couple d'amorce définit un locus précis sur le plasmide.

B. Réaction de PCR quantitative

Les ADN matrices sont préalablement incubés 5 min à 65°C puis conservés dans la glace. Dans un volume total de 6.04 µL, 2.8 µL de différentes dilutions d'ADN immunoprécipité, 0.24 µL de chaque paire d'amorce (200 nM) et 3 µL de Master MIX ABI SYBR Green 2X (réf 4367660) sont déposés dans chaque puits (plaque de 384 puits). Chaque mesure est effectuée en triplicata. La PCR quantitative est réalisée dans un thermocycleur (Viia 7 by life technologie) de 384 puits de chez Applied Biosystems®. Le cycle de PCR est le suivant : 10 min à 95°C puis 40 cycles à deux étapes (10 s à 95°C puis 30 s à 60°C). Les amorces sont validées de deux façons. Pour vérifier la pureté du produit PCR, une courbe de fusion est réalisée (de 60°C à 95°C en 20 minutes). La présence d'un seul pic autour de 81°C signifie que les amorces ne font pas de dimères d'amorce et qu'il n'y a pas de produit d'amplification secondaire. Pour s'assurer que les amorces peuvent servir à faire de la quantification, on calcule l'efficacité de la PCR avec l'équation $[E = (10^{(-1/\text{pente})} - 1) \times 100]$ après avoir réalisé une courbe standard à partir de dilutions d'ADN ("input" NIP). Les efficacités PCR de chaque couple d'amorces sont toutes supérieures à 95%. Pour pouvoir comparer les amplifications des couples entre eux, ils doivent avoir une efficacité PCR similaire.

C. Analyse des réactions de PCR quantitative

L'analyse des données brutes est effectuée à l'aide du programme ViiA 7 RUO Software ainsi que sur Excel. La quantification des loci immunoprécipités est relative au locus amplifié

par AS18s/as. L'échantillon d'"input" NIP est utilisé pour réaliser une gamme de dilution standard (dilution de 4 en 4) et permettre une quantification précise de l'ADN précipité par la ChIP (fraction IP) comparé au contrôle négatif traité sans anticorps anti-SopB (fraction IP(-)). Après amplification PCR, une valeur de C_T est attribuée à chaque échantillon. Pour réaliser la courbe standard, il faut exprimer les valeurs de C_T des différentes dilutions de l'input en fonction du $\log Q_0$ (Q_0 = quantité de départ). On obtient une droite dont la pente est proche de -3.32 lorsque l'efficacité PCR est de 100%. Grâce à cette courbe standard, on obtient une valeur de pente (a) et d'interception de la courbe avec l'axe des ordonnées (b). Pour réaliser la quantification des différents loci après IP, on utilise la formule $Q = 10^{((C_T - b)/a)}$, où C_T est la valeur assignée pour chaque locus. L'ensemble des quantités obtenues pour chaque locus est normalisé par rapport au locus du couple AS18.

XV. Séquençage haut débit

Les échantillons issus de ChIP sont ensuite resuspendus dans 10 mM de Tris-HCl, pH 8.5. Le séquençage haut débit a été réalisé avec le kit de séquençage (50 cycles Truseq SBS Kit v3 (FC-401-3002, Illumina)), sur un séquenceur (Hi-seq) de la plateforme d'IMAGIF (Centre de Recherche de Gif - www.imagif.cnrs.fr). L'analyse des données a été faite directement par la plateforme avec le logiciel CASAVA-1.8.2 pour le démultiplexage et BWA-0.6.2 pour les alignements. Ce travail a donc bénéficié des installations et de l'expertise de cette plateforme, notamment pour l'analyse bio-informatique qui en découle (cartographie sur génome de référence)

Tableau 6 | Plasmides utilisés dans cette étude

| Nom | Caractéristiques | Sources/références |
|-----------|--|-----------------------------------|
| pOX38-B | pOX38, <i>DccdB::cat</i> (gift from A. Wright) | (Guyer et al., 1981) |
| pZC204 | pACYC184, <i>DtetAWsopC+</i> , <i>cat+</i> | (Yates et al., 1999) |
| pZC302 | pBR322, <i>sopC</i> (43 pb) | (Bouet et al., 2005) |
| pBSKS(+) | | (Stratagene) |
| pAM238 | <i>lacOPZ'</i> Wmcs in <i>Eco RI-Hind III</i> interval of pGB2, <i>aadA+</i> | (Bouet et al., 1996) |
| pBBR1MCS5 | <i>rep</i> pBBR1, <i>mob</i> RK2, <i>sacRB</i> , <i>strAB</i> | (Kovach et al., 1995) |
| pYAS6 | <i>pTYB1</i> , <i>pT7::sopB-G324::Sce Intein::CBD</i> | (Ah-Seng et al., 2009) |
| pYAS47 | pAM238, <i>plac::sopAB</i> | Ah-Seng, unpubl. |
| pAS1 | pYAS6, <i>pT7::sopB-Δ23-G324::Sce Intein::CBD</i> | Ce travail |
| pAS2 | pYAS6, <i>pT7::sopB-Δ9-G324::Sce Intein::CBD</i> | Ce travail |
| pAS3 | pYAS6, <i>pT7::sopB-Δ8-G324::Sce Intein::CBD</i> | Ce travail |
| pAS4 | pYAS6, <i>pT7::sopB-4-G324::Sce Intein::CBD</i> | Ce travail |
| pAS5 | pYAS6, <i>pT7::sopB-5-G324::Sce Intein::CBD</i> | Ce travail |
| pAS6 | pYAS6, <i>pT7::sopB-6-G324::Sce Intein::CBD</i> | Ce travail |
| pAS7 | pYAS6, <i>pT7::sopB-7-G324::Sce Intein::CBD</i> | Ce travail |
| pAS8 | pYAS6, <i>pT7::sopB-8-G324::Sce Intein::CBD</i> | Ce travail |
| pAS9 | pYAS6, <i>pT7::sopB-9-G324::Sce Intein::CBD</i> | Ce travail |
| pAS10 | pYAS6, <i>pT7::sopB-10-G324::Sce Intein::CBD</i> | Ce travail |
| pAS11 | pYAS6, <i>pT7::sopB-11-G324::Sce Intein::CBD</i> | Ce travail |
| pAS12 | pYAS6, <i>pT7::sopB-12-G324::Sce Intein::CBD</i> | Ce travail |
| pAS13 | pYAS6, <i>pT7::ParB-G291::Sce Intein::CBD avec ParB de P. putida</i> | (Sanchez et al., 2013) |
| pAS14 | pYAS6, <i>pT7::ParB-R194A-G291::Sce Intein::CBD avec ParB de P. putida</i> | (Sanchez et al., 2013) |
| pAS15 | pYAS6, <i>pT7::ParB-G306::Sce Intein::CBD avec ParB de B. cenocepacia chr1</i> | (Sanchez et al., 2013) |
| pAS16 | pYAS6, <i>pT7::ParB-R192A-G306::Sce Intein::CBD avec ParB de B. cenocepacia chr1</i> | (Sanchez et al., 2013) |
| pAS17 | pDAG114, <i>sopB-11::mEos2</i> | Ce travail |
| pAS18 | pDAG114, <i>sopB-12::mEos2</i> | Ce travail |
| pAS19 | pDAG114, <i>sopB-AAA::R121-R122R123::mEos2</i> | Ce travail |
| pAS20 | pDAG114, <i>sopB-11::mVenus-Yfp</i> | Ce travail |
| pAS21 | pDAG114, <i>sopB-12::mVenus-Yfp</i> | Ce travail |
| pAS22 | pDAG114, <i>sopB-AAA::R121-R122R123::mVenus-Yfp</i> | Ce travail |
| pAS23 | pDAG170, <i>para_{BAD}::SopB-11</i> | Ce travail |
| pAS24 | pDAG170, <i>para_{BAD}::SopB-12</i> | Ce travail |
| pAS25 | pDAG114, <i>sopB-R108A</i> | Ce travail |
| pAS26 | pDAG114, <i>sopB-11</i> | Ce travail |
| pAS27 | pDAG114, <i>sopB-12</i> | Ce travail |
| pAS28 | pDAG114, <i>sopB-R108K</i> | Ce travail |
| pAS29 | pJYB67, <i>kn::Cm</i> | |
| pDAG114 | <i>repFIA+</i> , <i>ccdB-</i> , <i>resD+</i> , <i>rsfF+</i> , <i>cat+</i> | (Ravin and Lane, 1999) |
| pDAG127 | pFUS2, <i>SD-SopA⁺</i> inséré au site <i>EcoRI</i> | (Lemonnier et al., 2000) |
| pDAG170 | <i>para_{BAD}::sopB</i> | (Bouet et al., 2007) |
| pDAG209 | pDAG114, <i>DsopAB</i> | (Bouet et al., 2006) |
| pDAG502 | pBBR1MCS5, <i>with parAB from P. putida</i> | (Godfrin-Estevenson et al., 2002) |
| pDAG504 | pDAG114, <i>DsopABC parS+</i> | (Godfrin-Estevenson et al., 2002) |
| pDAG562 | pBBR1MCS5 <i>with parAB from B. cenocepacia chr1</i> | (Dubarry et al., 2006) |
| pJYB55 | pBSKS ⁺ , <i>bla,sopC⁺</i> (43pb) | (Bouet et al., 2006) |
| pJYB57 | pBSKS ⁺ , <i>bla,sopC⁺</i> | (Ah-Seng et al., 2009) |
| pJYB160 | pDAG127, <i>sop-GSKCK</i> | Ce travail |
| pJYB161 | pYAS47, <i>sopB-R219A</i> | (Sanchez et al., 2013) |
| pJYB162 | pYAS47, <i>sopB-R219K</i> | (Sanchez et al., 2013) |
| pJYB167 | pDAG562, <i>parB-R192A (Bcen)</i> | (Sanchez et al., 2013) |
| pJYB168 | pDAG502, <i>parB-R194A (Pput)</i> | (Sanchez et al., 2013) |
| pJYB208 | pYAS6, <i>pT7::sopB-R219A-G324::Sce Intein::CBD</i> | (Sanchez et al., 2013) |
| pJYB209 | pYAS6, <i>pT7::sopB-R219K-G324::Sce Intein::CBD</i> | (Sanchez et al., 2013) |
| pJYB212 | pDAG114, <i>sopB::mEos2</i> | Ce travail |
| pJYB214 | pDAG114, <i>sopB::Venus-Yfp</i> | (Sanchez et al., 2013) |
| pJYB216 | pJYB212, <i>DsopC</i> | Ce travail |
| pJYB222 | pJY212, <i>sopB-R219A::mEos2</i> | Ce travail |
| pJYB223 | pJY214, <i>sopB-R219A::Venus-Yfp</i> | (Sanchez et al., 2013) |
| pJYB234 | pDAG214, <i>sopB::mVenus-Yfp</i> | Ce travail |
| pJYB243 | pDAG114, <i>sopA::mVenus-Yfp</i> | Ce travail |
| pJYB226 | pJY212, <i>sopB-A183D-K191A-R195A::mEos2</i> | (Sanchez et al., 2013) |

Tableau 7 | Oligonucléotides utilisés pour l'expérience de gel EMSA et de Footprint

| Nom des oligonucléotides | | Séquence 5' -> 3' | Matrice | Détails |
|--------------------------|--|-------------------|-----------------------|---|
| sopC 22 | CTAGTGGATCCCCGGGCTGCAG | | pJYB55 | sonde <i>sopC144</i> |
| sopC 23 | CCCTCACTAAAGGGAACAAAAGC | | | |
| sopC 22bis | CTAGTGGATCCCCCG | | pJYB55 | sonde <i>sopC144</i> |
| sopC 23bis | CCCTCACTAAAGGGA | | | |
| sopC34 | CAGTGAATTGTAATACGACTCAC | | | |
| sopC164a | TCGGAATTAACCTCACTAA | | pJYB55, pJYB58 | sonde <i>sopC223</i> , sonde <i>sopC266</i> |
| repE 1 | AGTTACCAGCGTATGCCGACTTCC | | | |
| repE 2 | ACGATATGAGTCGCTGGCGGCCTT | | pDAG114 | sonde <i>repE114</i> |
| sopC 37t | ATTAGTCGCACGTGCAACGCTCTCGTAT | | | |
| sopC 37b | ATACGAGGACGCGTTGCACGTGCGACTAAT | | Hybridation entre eux | Sonde <i>sopC37</i> |
| sopC 38t | ATTAGTCTGGGACACGGTCCCAGCTCGTAT | | | |
| sopC 38b | ATACGAGTGGGACCGTGGTCCCAGACTAAT | | Hybridation entre eux | Sonde <i>sopC38</i> |
| sopC 39t | ATTAGTCCGAGACACGGTCTCGCTCGTAT | | | |
| sopC 39b | ATACGAGCGAGACCGTGGTCTCGGACTAAT | | Hybridation entre eux | Sonde <i>sopC39</i> |
| sopC 40t | ATTAGTCCGAAACACGCTTTCGCTCGTAT | | | |
| sopC 40b | ATACGAGCGAAACCGTGGTTCGACTAAT | | Hybridation entre eux | Sonde <i>sopC40</i> |
| sopC 41t | ATTAGTCCGAACCAACGGTTCGCTCGTAT | | | |
| sopC 41b | ATACGAGCGAACCCGTGGTTCGACTAAT | | Hybridation entre eux | Sonde <i>sopC41</i> |
| sopC 42t | ATTAGTCTGGGACACACGCGTCTCGTAT | | | |
| sopC 42b | ATACGAGGACGCGTGTGTCAGACTAAT | | Hybridation entre eux | Sonde <i>sopC42</i> |
| sopC 43t | ATTAGTCCGGGACACGGTCCCCTCGTAT | | | |
| sopC 43b | ATACGAGCGGGACCGTGGTCCGGACTATT | | Hybridation entre eux | Sonde <i>sopC43</i> |
| sopC 44t | ATTAGTCTGAGACACGGTCTCACTCGTAT | | | |
| sopC 44b | ATACGAGTGAGACCGTGGTCTCAGACTATT | | Hybridation entre eux | Sonde <i>sopC44</i> |
| sopC 45t | ATTAGTCTGGGACACGGTCCCAGACTAAT AATCAGGTGCG | | | |
| sopC 45b | CGACCTGATTATTAGTCTGGGACCGTGGTC CCAGACTAAT | | Hybridation entre eux | Sonde <i>sopC45</i> |
| sopC 46t | ATTAGTCTGGGACACCGTCCCAGACTAAT AATCAGGTGCG ACCTCGAGGG GGGGCCCGGT | | | |
| sopC 46b | ACCGGGCCCCCTCGAGGTGACCTGATT ATTAGTCTGG GACCGTGGTCCCAGACTAAT | | Hybridation entre eux | Sonde <i>sopC46</i> |
| sopC 47t | TCTGGGACACCGTCCCAGA | | | |
| sopC 47b | TCTGGGACCGTGGTCCCAGA | | Hybridation entre eux | Sonde <i>sopC47</i> |
| sopC 48t | ATTAGTCTGGGACACCGTCCCAGCTCGTATCGTCGACGGT | | | |
| sopC 48b | ACCGTCGACGATACGAGTGGGACCGTGGTCCCAGACTAAT | | Hybridation entre eux | Sonde <i>sopC48</i> |
| sopC 49t | ATTAGTCTGGGACACCGTCCCAGCTCGTATCGTCGACGGTATCGATAAGCTTGATATCGA | | | |
| sopC 49b | TCGATATCAAGCTTATCGATACCGTCGACGATACGAGTGGGACCGTGGTCCCAGACTAAT | | Hybridation entre eux | Sonde <i>sopC49</i> |
| sopC 50t | GAGGTGCACTGATTATTAGTCTGGGACACGGTCCCAGCTCGTATGTCGACGGTATCGA | | | |
| sopC 50b | TCGATACCGTCTGACGATACGAGTGGGACCGTGGTCCCAGACTAATAATCAGGTGACACTC | | Hybridation entre eux | Sonde <i>sopC50</i> |
| sopC 51t | TGATTATTAGTCTGGGACACGGTCCCAGCTCGTATCGTCG | | | |
| sopC 51b | CGACGATACGAGTGGGACCGTGGTCCCAGACTAATAATCA | | Hybridation entre eux | Sonde <i>sopC51</i> |
| sopC 52t | ATTATTAGTCTGGGACACGGTCCCAGCTCGTATCGT | | | |
| sopC 52b | ACGATACGAGTGGGACCGTGGTCCCAGACTAATAAT | | Hybridation entre eux | Sonde <i>sopC52</i> |
| sopC 53t | TCTGGGACACCGTCCCAGCTGATCGT | | | |
| sopC 53b | ACGATACGAGTGGGACCGTGGTCCCAGA | | Hybridation entre eux | Sonde <i>sopC53</i> |
| sopC 54t | TCTGGGACACCGTCCCAGCTGATCGTCG | | | |
| sopC 54b | CGACGATACGAGTGGGACCGTGGTCCCAGA | | Hybridation entre eux | Sonde <i>sopC54</i> |
| sopC 55t | TCTGGGACACCGTCCCAGCTCGTATCGTCGAC | | | |
| sopC 55b | GTCGACGATACGAGTGGGACCGTGGTCCCAGA | | Hybridation entre eux | Sonde <i>sopC55</i> |
| sopC 56t | TCTGGGACACCGTCCCAGCTCGTATCGTCGACGGTATCGA | | | |
| sopC 56b | TCGATACCGTCTGACGATACGAGTGGGACCGTGGTCCCAGA | | Hybridation entre eux | Sonde <i>sopC56</i> |
| sopC 57t | TCTGGGACACCGTCCCAGCTCGTATCGTCGACGGTATCGATAAGCTTGATATCGAATCTCTGC | | | |
| sopC 57b | GCAGGAATTTCGATATCAAGCTTATCGATACCGTCGACGATACGAGTGGGACCGTGGTCCCAGA | | Hybridation entre eux | Sonde <i>sopC57</i> |
| sopC 58t | ACTGGGACACCGTCCCAGCTCGTATCGTCGGTCTGATTATTAGTCTGGGACACGGTCCCAGT | | | |
| sopC 58b | AGTGGGACCGTGGTCCCAGACTAATAATCAGACCGACGATACGAGTGGGACCGTGGTCCCAGT | | Hybridation entre eux | Sonde <i>sopC58</i> |
| sopC 59t | ACTGGGACACCGTCCCAGCTCGTATCGTCGGTCTGATTATTGGGACACGGTCCCAGCTCGTAT | | | |
| sopC 59b | ATACGAGTGGGACCGTGGTCCCAGTATCAGACCGACGATACGAGTGGGACCGTGGTCCCAGT | | Hybridation entre eux | Sonde <i>sopC59</i> |
| sopC 60t | ACTGGGACACCGTCCCAGCTCGTATCGTCGGTCTTGGGACACGGTCCCAGTATCGTCGG | | | |
| sopC 60b | CCGACGATACGAGTGGGACCGTGGTCCCAGACGACGATACGAGTGGGACCGTGGTCCCAGT | | Hybridation entre eux | Sonde <i>sopC60</i> |
| sopC 61t | ACTGGGACACCGTCCCAGCTCGTATCGTCTGGGACACGGTCCCAGTATCGTCGGTCTGA | | | |
| sopC 61b | TCAGACCGACGATACGAGTGGGACCGTGGTCCCAGACGATACGAGTGGGACCGTGGTCCCAGT | | Hybridation entre eux | Sonde <i>sopC61</i> |
| sopC 62t | ACTGGGACACCGTCCCAGTCTGGGACACCGTCCCAGTATCGTCGGTCTGATTATTA | | | |
| sopC 62b | TAATAATCAGACCGACGATACGAGTGGGACCGTGGTCCCAGAGTGGGACCGTGGTCCCAGT | | Hybridation entre eux | Sonde <i>sopC62</i> |
| sopC63t(E15t) | ATTAGTCTGGGACAACTGTCCCAGCTCGTAT | | | |
| sopC63b(E15b) | ATACAGAGTGGGACAGTTGTCCCAGACTAAT | | Hybridation entre eux | Sonde <i>sopC63</i> |
| sopC64t(E64t) | ATTAGTCTGGGACTACAGTCCCAGCTCGTAT | | | |
| sopC64b(E64b) | ATACGAGTGGGACTGATGCCAGACTAAT | | Hybridation entre eux | Sonde <i>sopC64</i> |
| sopC65t | ACTGGGACACCGTCCCAGTGGGACACCGTCCCAGTATCGTCGGTCTGATTATTAGTCTCT | | | |
| sopC65b | AGGACTAATAATCAGACCGACGATACGAGTGGGACCGTGGTCCCAGTGGGACCGTGGTCCCAGT | | Hybridation entre eux | Sonde <i>sopC65</i> |
| sopC66t | ATTAGTCTGTCCACGTGGAACACTCGTAT | | | |
| sopC66b | ATACGAGTGTTCACGTGGAACAGACTAAT | | Hybridation entre eux | Sonde <i>sopC66</i> |
| sopC67t | ATTAGTCTGTTCACGTGAAACACTCGTAT | | | |
| sopC67b | ATACGAGTGTTCACGTGAAACAGACTAAT | | Hybridation entre eux | Sonde <i>sopC67</i> |

Tableau 8 | Oligonucléotides utilisés pour l'expérience de qChIP

| Nom des oligonucléotides | Séquence 5' -> 3' | Matrice |
|--------------------------|-----------------------|--------------------|
| AS18 L | GAAAAGCCAGCACCTGATG | pDAG114, pOX38-B |
| AS18 R | GAACAGTGGGCCAGAGAGA | |
| AS19 L | TGGGAACCTGTCTGCAATGA | |
| AS19 R | CTCCTACGCGCGCAATTAAC | |
| AS20 L | GGCGCTGGAGAATAGGTGAA | |
| AS20 R | ACCCGTCCTCGAATTTCCAT | |
| AS21 L | CGGCGATGAAAAAGGCTATG | |
| AS21 R | CATGGCATA CGGATTGGTGA | |
| AS22 L | TCCATCAATGCCTGTCAAGG | pDAG114 |
| AS22 R | ACAGCAGCAATCCACAGCAG | |
| AS23 L | TAAGCGGGCAAAGTCGGTTA | |
| AS23 R | GAGGACGCCGAATTTTAGGG | |
| AS24 L | AGGCGTTTAAGGGACCAAT | |
| AS24 R | CAGGTTTCATGCGCGTTTG | |
| AS 35 L | AACGGTCAGGACCTGGATTG | |
| AS 35 R | AGCATGTGCATCGCATAGGA | |
| AS26 L | GGCCTTGTGGGTATGAGGTG | pOX38-B |
| AS26 R | ACGGCGGCTTAATACTGCAA | |
| AS27 L | CGGAGGCTGCGTAAATGAAC | |
| AS27 R | ATCGGAATCAAAGCGACAGC | |
| AS28 L | GCGACCGTTTTTCATCTGGTC | |
| AS28 R | CGTGGTACAGCCACAACCTGG | |
| AS29 L | ACTGGTTCCCTCAGCACAG | |
| AS29 R | GTGCAGTTCTTCTGCGGTCA | |
| AS30 L | GCACTGAAACTGCCTGACGA | |
| AS30 R | AAGCCCATCGTGGTGTCATT | |
| AS 36 L | TTGAGAAGCGTCGACAGGAA | |
| AS 36 R | ACACGAAGCCCCAGTTCAAG | |
| AS 37 L | ATTGACTGCATGGCGAAGGT | |
| AS 37 R | ACCACGCTGGACACACACAT | |
| AS 38 L | TACGTTCCGAAACCTGTCA | |
| AS 38 R | TGGGAACATGGACTTGCTCA | |
| AS31 L | AATTGGCTATCCGCGTGTGT | pDAG209 |
| AS31 R | CTGGCGAGGAAGCAAAGAAG | |
| AS 34 L | TCCGAGCTTGGCGTAATCAT | pJYB 57 |
| AS 34 R | GCTTCCGGCTCGTATGTTGT | |
| AS 42 L | ACATGGCCCGCTTTATCAGA | pZC302 |
| AS 42 R | GCAGCTGCGGTAAAGCTCAT | |
| dnaA1 L | ATGCGAGCTTAACCACTCA | Chromosome DLT1215 |
| dnaA1 R | CGTGCTCGTATGGCTACTA | |
| terC2 L | CCACCAATCCGTTGATGTAA | |
| terC2 R | TAAGGAGACCCTGTATGGC | |

Tableau 9 | Oligonucléotides utilisés pour la construction des plasmides de cette étude

| Nom des oligonucléotides | | Séquence 5' -> 3' | Matrice | Détails |
|--------------------------|--|-------------------|---------------------------|----------------------------|
| SopB-AS1 | CGGGAGCGGATTTGAACGTTGC | | pYAS6 | Construction des délétions |
| SopB-AS2 | AGCACTTGTTGGGCATAATCG | | | |
| SopB-AS-D1 | TACATATGAAGCGTGCCCCGATGGTGATTGTTAATTG | | pYAS6 | Construction sopBΔ23 |
| SopB-AS-D2 | TCCACCATCGGGGCACGCTTCATATGATATCTCTTC | | | |
| SopB-AS-D3 | CGCTCAATACTCAACCAGCTGCCCCGATGGTGATTGCG | | pYAS6 | Construction sopBΔ9 |
| SopB-AS-D4 | CCATCGGGGCAGCTGTTGAGTATTGAGCGTATGTTTTGG | | | |
| SopB-AS-D5 | GTGCGCCTGTTATTCGGTTGAAGATACTTCGTTATCGAC | | pYAS6 | Construction sopBΔ8 |
| SopB-AS-D6 | GTATCTCAACCGGAATAACAGGCGCACGCTTCATATG | | | |
| SopB-D43/76-L | GTAATGGCTCGCGGTAATGAACGTGACCGAGGAGCTGCTT | | pYAS6 | Construction sopB-4 |
| SopB-D43/76-R | GTCACGTTCAATACCGGAGCCATTACTCTACGCGC | | | |
| SopB-D43/64-L | GTAATGGCTCGCGGTAGATGTTGAGAAGACCTCTC | | pYAS6 | Construction sopB-5 |
| SopB-D43/64-R | AACACTATCACCGGAGCATTACTCTACGCGGCAAT | | | |
| SopB-D64/76-L | GAAGTGCTCCGGGTAATGAACGTGACCGAGGAGCTGCTT | | pYAS6 | Construction sopB-6 |
| SopB-D64/76-R | GTCACGTTCAATACCGGAGCACTTCAAGAGTAACTTC | | | |
| SopB-D50/73-L | CATTACTTTGCTGTATGGTCAGGTAATGAACGTGAC | | pYAS6 | Construction sopB-7 |
| SopB-D50/73-R | ACCTGACCATACAGGCAAGTAATGGCATTACCGGAGC | | | |
| AS16 R | AGAAGGAGATATACATATGAAG | | pYAS6, pAS8 | pAS11, pAS12 |
| AS16 L | CTTCTCAACACTATCACCCGGAAGC | | | |
| AS17 R | GATAGTGTGAGAAGGCTGCCGCGTAGCGGCAGCTAATGAACGTGACCAG | | pBBABK1. | pJYB167 |
| AS17 L | GTGAGCTTAACTAGTTCTTG | | | |
| AS 32 L | CATGGGGCACGCAgcaGCGCTGCTCGCGGTC | | pBBABK1. | pJYB167 |
| AS 32 R | ACCGCGAGCAGCGCtgctGCGTGCCCCATG | | | |
| AS 32 L-screen | CGACATGGGGCACGCAgca | | pDAG502 | pJYB168 |
| AS 32 R-screen | GCAACTGTTGGGAAGGCG | | | |
| AS 33 L | GAGATGGGGCACGCGcaGCAATTGCTGGGCCTG | | pDAG502 | pJYB168 |
| AS 33 R | CAGGCCCAAGCAATGctgGCGGTGCCCATCTC | | | |
| AS 33 L-screen | gGAGATGGGGCACGCGca | | pAS11, pAS12 | pAS 17, pAS18 |
| AS 33 R-screen | AAGGGTCAACACGAAGAGC | | | |
| AS 44 | ATGAAATTTTCGATCGTCTGATTAAACACGCTGGGAGATTAGATAATGAAGCGTGCGCTGTTATTC | | pDAG 503, pJYB 168 (pput) | pAS 13, pAS14 |
| AS 45 L | AGAAGGAGATATACATATGGCCGTCAAGAAACGGGTCTC | | | |
| AS 45 R | CATTAAACATTGTGACCTTGCAAGCACCCGCGATGTGAGCGAGCACGCC | | pDAG 562, pJYB 167 (Bcen) | pAS 15, pAS16 |
| AS 46 L | AGAAGGAGATATACATATGAGCGAAGCGAGGAAGAAAGAC | | | |
| AS 46 R | CATTAAACATTGTGACCTTGCAAGCACCCCTCTTCGGTGCACGTTGCC | | pJYB 67 | pAS29 |
| AS47 L | TTTAGCTTCCTTAGCTCCTG | | | |
| AS47 R | TTTTTTAAGGCAGTTATTGGTGCCC | | | |

Tableau 10 *Œ Souches utilisées dans cette étude*

| Nom | Génotype | Sources/références |
|-----------|---|----------------------------|
| DLT 1913 | thi, leu, thyA, deoB, supE, Δ(ara-leu)7679, zac3051::Tn10, rpsL::pcp18 | |
| DLT 2704 | DLT 1913, pJYB160 | |
| JS238 | araD139, Δ(ara-leu)7679, ΔlacX74, galU, galK, rpsL, thi, HsdR2, mcrB, P _{malE} ::lacI ^Q , srlC::Tn10, recA1 | |
| DLT 2317 | JS238, pYAS6 | Ce travail |
| DLT 2630 | JS238, pAS1 | Ce travail |
| DLT 2631 | JS238, pAS2 | Ce travail |
| DLT 2632 | JS238/pAS3 | Ce travail |
| DLT1215 | F ⁻ , thi, leu, thyA, deoB, supE, Δ(ara-leu)7696, zac3051::Tn10, rpsL812 | (Bouet et al., 2006) |
| DLT 1780 | DLT 1215, pDAG114 | |
| DLT2888 | DLT 1215, pOX38-B conjugation from NCM3722/pOX38-B | Ce travail |
| DLT 2986 | DLT 1215, pDAG114, pJYB57 | Ce travail |
| DLT 2986 | DLT 1215, pOX38B, pJYB57 | Ce travail |
| DLT 2999 | DLT 1215, pDAG114, pZC302 | Ce travail |
| DLT 3000 | DLT 1215, pOX38B, pJYB60 | Ce travail |
| DLT 3001 | DLT 1215, pDAG114, pJYB67 | Ce travail |
| DLT 3002 | DLT 1215, pOX38B, pJYB67 | Ce travail |
| DLT 3143 | DLT 1215, pAS23 | Ce travail |
| DLT 3040 | DLT 1215, pAS24 | Ce travail |
| DLT 3144 | DLT 1215, pAS25 | Ce travail |
| DLT 3145 | DLT 1215, pAS26 | Ce travail |
| DLT 3041 | DLT 1215, pAS27 | Ce travail |
| DLT 3042 | DLT 1215, pAS28 | Ce travail |
| DLT 3043 | DLT 1215, pAS29 | Ce travail |
| DLT 3044 | DLT 1215, pOX38B, pAS29 | Ce travail |
| C2833 | F ⁺ , proA+B+, lacIqzrf::Tn10(TetR)/ thuA2, ompT, lon, dcm, gal, lacZ::gene1 T7, sulA11, endA1, Δ(mcr-73::mTn10,TcS), Δ(γgb-210::Tn10, TcS), Δ(mcrC- | NEBioLabs |
| DLT 2652 | C2833, pAS1 | Ce travail |
| DLT 2653 | C2833, pAS2 | Ce travail |
| DLT 2654 | C2833, pAS3 | Ce travail |
| DLT 3146 | C2833, pAS8 | Ce travail |
| DLT 3147 | C2833, pAS9 | Ce travail |
| DLT 3148 | C2833, pAS10 | Ce travail |
| DLT 3149 | C2833, pAS11 | Ce travail |
| DLT 3150 | C2833, pAS12 | Ce travail |
| DLT 3013 | C2833, pAS13 | Ce travail |
| DLT 3014 | C2833, pAS14 | Ce travail |
| DLT 3015 | C2833, pAS15 | Ce travail |
| DLT 3016 | C2833, pAS16 | Ce travail |
| BL21(DE3) | F ⁻ omp T hsd SB(rB ⁻ , mB ⁻) gal dcm lon- clpP ⁻ (DE3) | Equipe de P. Genevaux LMGM |
| DLT 3151 | BL21(DE3), pYAS6 | Ce travail |
| DLT 3152 | BL21(DE3), pAS4 | Ce travail |
| DLT 3153 | BL21(DE3), pAS5 | Ce travail |
| DLT 3154 | BL21(DE3), pAS6 | Ce travail |
| DLT 3155 | BL21(DE3), pAS7 | Ce travail |
| Stellar | F ⁻ , endA1, supE44, thi-1, recA1, relA1, gyrA96, phoA, Φ80d lacZA M15, Δ(lacZYA - argF) U169, Δ (mrr - hsdRMS - mcrBC), ΔmcrA, λ- | Clontech |
| DLT 3156 | Stellar, pAS17 | Ce travail |
| DLT 3157 | Stellar, pAS18 | Ce travail |
| DLT 3158 | Stellar, pAS19 | Ce travail |
| DLT 3159 | Stellar, pAS20 | Ce travail |
| DLT 3160 | Stellar, pAS21 | Ce travail |
| DLT 3161 | Stellar, pAS22 | Ce travail |

Bibliographie

- Abeles, A.L., Friedman, S.A., and Austin, S.J. (1985) Partition of unit-copy miniplasmids to daughter cells. III. The DNA sequence and functional organization of the P1 partition region. *J Mol Biol* **185**: 261–272.
- Adachi, S., Hori, K., and Hiraga, S. (2006) Subcellular positioning of F plasmid mediated by dynamic localization of SopA and SopB. *J Mol Biol* **356**: 850–863.
- Ausmees, N., Wahlstedt, H., Bagchi, S., Elliot, M.A., Buttner, M.J., and Flärdh, K. (2007) SmeA, a small membrane protein with multiple functions in *Streptomyces* sporulation including targeting of a SpoIIIE/FtsK-like protein to cell division septa. *Mol Microbiol* **65**: 1458–1473.
- Austin, S., and Abeles, A. (1983) Partition of unit-copy miniplasmids to daughter cells. I. P1 and F miniplasmids contain discrete, interchangeable sequences sufficient to promote equipartition. *J Mol Biol* **169**: 353–372.
- Austin, S., and Nordström, K. (1990) Partition-mediated incompatibility of bacterial plasmids. *Cell* **60**: 351–354.
- Austin, S., and Wierzbicki, A. (1983) Two mini-F-encoded proteins are essential for equipartition. *Plasmid* **10**: 73–81.
- Austin, S., Ziese, M., and Sternberg, N. (1981) A novel role for site-specific recombination in maintenance of bacterial replicons. *Cell* **25**: 729–736.
- Aylett, C.H., and Löwe, J. (2012) Superstructure of the centromeric complex of TubZRC plasmid partitioning systems. *Proc Natl Acad Sci* **109**: 16522–16527.
- Barillà, D., Carmelo, E., and Hayes, F. (2007) The tail of the ParG DNA segregation protein remodels ParF polymers and enhances ATP hydrolysis via an arginine finger-like motif. *Proc Natl Acad Sci U S A* **104**: 1811–1816.
- Barillà, D., and Hayes, F. (2003) Architecture of the ParF*ParG protein complex involved in prokaryotic DNA segregation. *Mol Microbiol* **49**: 487–499.
- Barillà, D., Rosenberg, M.F., Nobbmann, U., and Hayes, F. (2005) Bacterial DNA segregation dynamics mediated by the polymerizing protein ParF. *EMBO J* **24**: 1453–1464.
- Bartosik, A.A., Lasocki, K., Mierzejewska, J., Thomas, C.M., and Jagura-Burdzy, G. (2004) ParB of *Pseudomonas aeruginosa*: interactions with its partner ParA and its target parS and specific effects on bacterial growth. *J Bacteriol* **186**: 6983–6998.
- Bartosik, A.A., Mierzejewska, J., Thomas, C.M., and Jagura-Burdzy, G. (2009) ParB deficiency in *Pseudomonas aeruginosa* destabilizes the partner protein ParA and affects a variety of physiological parameters. *Microbiol Read Engl* **155**: 1080–1092.
- Becker, E., Herrera, N.C., Gunderson, F.Q., Derman, A.I., Dance, A.L., Sims, J., *et al.* (2006) DNA segregation by the bacterial actin AlfA during *Bacillus subtilis* growth and development. *EMBO J* **25**: 5919–5931.

Bergquist, P.L., Downard, R.A., Caughey, P.A., Gardner, R.C., and Lane, H.E. (1981) Analysis of mini-F plasmid replication by transposition mutagenesis. *J Bacteriol* **147**: 888–899.

Berlatzky, I.A., Rouvinski, A., and Ben-Yehuda, S. (2008) Spatial organization of a replicating bacterial chromosome. *Proc Natl Acad Sci* **105**: 14136–14140.

Biek, D.P., and Shi, J. (1994) A single 43-bp sopC repeat of plasmid mini-F is sufficient to allow assembly of a functional nucleoprotein partition complex. *Proc Natl Acad Sci U S A* **91**: 8027–8031.

Biek, D.P., and Strings, J. (1995) Partition functions of mini-F affect plasmid DNA topology in *Escherichia coli*. *J Mol Biol* **246**: 388–400.

Bignell, C.R., Haines, A.S., Khare, D., and Thomas, C.M. (1999) Effect of growth rate and incC mutation on symmetric plasmid distribution by the IncP-1 partitioning apparatus. *Mol Microbiol* **34**: 205–216.

Bigot, S., Sivanathan, V., Possoz, C., Barre, F.-X., and Cornet, F. (2007) FtsK, a literate chromosome segregation machine. *Mol Microbiol* **64**: 1434–1441.

Bingle, L.E.H., Macartney, D.P., Fantozzi, A., Manzoor, S.E., and Thomas, C.M. (2005) Flexibility in Repression and Cooperativity by KorB of Broad Host Range IncP-1 Plasmid RK2. *J Mol Biol* **349**: 302–316.

Blattner, F.R. (1997) The Complete Genome Sequence of *Escherichia coli* K-12. *Science* **277**: 1453–1462.

Blomberg, P., Nordström, K., and Wagner, E.G. (1992) Replication control of plasmid R1: RepA synthesis is regulated by CopA RNA through inhibition of leader peptide translation. *EMBO J* **11**: 2675.

Bouet, J.-Y., Bouvier, M., and Lane, D. (2006) Concerted action of plasmid maintenance functions: partition complexes create a requirement for dimer resolution. *Mol Microbiol* **62**: 1447–1459.

Bouet, J.Y., and Funnell, B.E. (1999) P1 ParA interacts with the P1 partition complex at parS and an ATP–ADP switch controls ParA activities. *EMBO J* **18**: 1415–1424.

Bouet, J.-Y., and Lane, D. (2009) Molecular basis of the supercoil deficit induced by the mini-F plasmid partition complex. *J Biol Chem* **284**: 165–173.

Bouet, J.Y., Nordström, K., and Lane, D. (2007) Plasmid partition and incompatibility—the focus shifts. *Mol Microbiol* **65**: 1405–1414.

Bouet, J.Y., Rech, J., Egloff, S., Biek, D.P., and Lane, D. (2005) Probing plasmid partition with centromere-based incompatibility. *Mol Microbiol* **55**: 511–525.

Bouet, J.Y., Ah-Seng, Y., Benmeradi, N., and Lane, D. (2007) Polymerization of SopA partition ATPase: regulation by DNA binding and SopB. *Mol Microbiol* **63**: 468–481.

- Bowman, G.R., Comolli, L.R., Zhu, J., Eckart, M., Koenig, M., Downing, K.H., *et al.* (2008) A polymeric protein anchors the chromosomal origin/ParB complex at a bacterial cell pole. *Cell* **134**: 945–955.
- Breier, A.M., and Grossman, A.D. (2007) Whole-genome analysis of the chromosome partitioning and sporulation protein Spo0J (ParB) reveals spreading and origin-distal sites on the *Bacillus subtilis* chromosome. *Mol Microbiol* **64**: 703–718.
- Breüner, A., Jensen, R.B., Dam, M., Pedersen, S., and Gerdes, K. (1996) The centromere-like parC locus of plasmid R1. *Mol Microbiol* **20**: 581–592.
- Britton, R.A., Lin, D.C., and Grossman, A.D. (1998) Characterization of a prokaryotic SMC protein involved in chromosome partitioning. *Genes Dev* **12**: 1254–1259.
- Brown, P.J.B., Hardy, G.G., Trimble, M.J., and Brun, Y.V. (2009) Complex regulatory pathways coordinate cell-cycle progression and development in *Caulobacter crescentus*. *Adv Microb Physiol* **54**: 1–101.
- Butan, C., Hartnell, L.M., Fenton, A.K., Bliss, D., Sockett, R.E., Subramaniam, S., and Milne, J.L.S. (2010) Spiral Architecture of the Nucleoid in *Bdellovibrio bacteriovorus*. *J Bacteriol* **193**: 1341–1350.
- Campbell, C.S., and Mullins, R.D. (2007) In vivo visualization of type II plasmid segregation: bacterial actin filaments pushing plasmids. *J Cell Biol* **179**: 1059–1066.
- Carmelo, E., Barillà, D., Golovanov, A.P., Lian, L.-Y., Derome, A., and Hayes, F. (2005) The unstructured N-terminal tail of ParG modulates assembly of a quaternary nucleoprotein complex in transcription repression. *J Biol Chem* **280**: 28683–28691.
- Castaing, J.-P., Bouet, J.-Y., and Lane, D. (2008) F plasmid partition depends on interaction of SopA with non-specific DNA. *Mol Microbiol* **70**: 1000–1011.
- Chaconas, G., and Kobryn, K. (2010) Structure, Function, and Evolution of Linear Replicons in *Borrelia*. *Annu Rev Microbiol* **64**: 185–202.
- Chattoraj, D.K. (2000) Control of plasmid DNA replication by iterons: no longer paradoxical. *Mol Microbiol* **37**: 467–476.
- Chen, C.W., Huang, C.-H., Lee, H.-H., Tsai, H.-H., and Kirby, R. (2002) Streptomyces chromosomes. *TRENDS Genet* **18**.
- Chen, Y., and Erickson, H.P. (2008) In vitro assembly studies of FtsZ/tubulin-like proteins (TubZ) from *Bacillus* plasmids: evidence for a capping mechanism. *J Biol Chem* **283**: 8102–8109.
- Choi, C.L., Claridge, S.A., Garner, E.C., Alivisatos, A.P., and Mullins, R.D. (2008) Protein-nanocrystal conjugates support a single filament polymerization model in R1 plasmid segregation. *J Biol Chem* **283**: 28081–28086.

-
- Cooper, S., and Helmstetter, C.E. (1968) Chromosome replication and the division cycle of *Escherichia coli* B/r. *J Mol Biol* **31**: 519–540.
- Cooper, V.S., Vohr, S.H., Wrocklage, S.C., and Hatcher, P.J. (2010) Why genes evolve faster on secondary chromosomes in bacteria. *PLoS Comput Biol* **6**: e1000732.
- Cortez, D., Quevillon-Cheruel, S., Gribaldo, S., Desnoues, N., Sezonov, G., Forterre, P., and Serre, M.-C. (2010) Evidence for a Xer/dif System for Chromosome Resolution in Archaea. *PLoS Genet* **6**: e1001166.
- Couturier, E., and Rocha, E.P.C. (2006) Replication-associated gene dosage effects shape the genomes of fast-growing bacteria but only for transcription and translation genes. *Mol Microbiol* **59**: 1506–1518.
- Cui, T., Moro-oka, N., Ohsumi, K., Kodama, K., Ohshima, T., Ogasawara, N., *et al.* (2007) *Escherichia coli* with a linear genome. *EMBO Rep* **8**: 181–187.
- Dabrazhynetskaya, A., Sergueev, K., and Austin, S. (2005) Species and incompatibility determination within the P1par family of plasmid partition elements. *J Bacteriol* **187**: 5977–5983.
- Dam, M., and Gerdes, K. (1994) Partitioning of plasmid R1. Ten direct repeats flanking the parA promoter constitute a centromere-like partition site parC, that expresses incompatibility. *J Mol Biol* **236**: 1289–1298.
- Dame, R.T., Luijsterburg, M.S., Krin, E., Bertin, P.N., Wagner, R., and Wuite, G.J.L. (2005) DNA Bridging: a Property Shared among H-NS-Like Proteins. *J Bacteriol* **187**: 1845–1848.
- Danilova, O., Reyes-Lamothe, R., Pinskaya, M., Sherratt, D., and Possoz, C. (2007) MukB colocalizes with the oriC region and is required for organization of the two *Escherichia coli* chromosome arms into separate cell halves. *Mol Microbiol* **65**: 1485–1492.
- Davey, M.J. (1997) Modulation of the P1 Plasmid Partition Protein ParA by ATP, ADP, and P1 ParB. *J Biol Chem* **272**: 15286–15292.
- Davey, M.J., and Funnell, B.E. (1994) The P1 plasmid partition protein ParA. A role for ATP in site-specific DNA binding. *J Biol Chem* **269**: 29908–29913.
- Davis, M.A., Martin, K.A., and Austin, S.J. (1990) Specificity switching of the P1 plasmid centromere-like site. *EMBO J* **9**: 991.
- Davis, M.A., Martin, K.A., and Austin, S.J. (1992) Biochemical activities of the parA partition protein of the P1 plasmid. *Mol Microbiol* **6**: 1141–1147.
- Davis, M.A., Radnedge, L., Martin, K.A., Hayes, F., Youngren, B., and Austin, S.J. (1996) The P1 ParA protein and its ATPase activity play a direct role in the segregation of plasmid copies to daughter cells. *Mol Microbiol* **21**: 1029–1036.

- Davis, T.L., Helinski, D.R., and Roberts, R.C. (1992) Transcription and autoregulation of the stabilizing functions of broad-host-range plasmid RK2 in *Escherichia coli*, *Agrobacterium tumefaciens* and *Pseudomonas aeruginosa*. *Mol Microbiol* **6**: 1981–1994.
- Delbrück, H., Ziegelin, G., Lanka, E., and Heinemann, U. (2002) An Src homology 3-like domain is responsible for dimerization of the repressor protein KorB encoded by the promiscuous IncP plasmid RP4. *J Biol Chem* **277**: 4191–4198.
- Derman, A.I., Lim-Fong, G., and Pogliano, J. (2008) Intracellular mobility of plasmid DNA is limited by the ParA family of partitioning systems. *Mol Microbiol* **67**: 935–946.
- Dillon, S.C., and Dorman, C.J. (2010) Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nat Rev Microbiol* **8**: 185–195.
- Disqué-Kochem, C., and Eichenlaub, R. (1993) Purification and DNA binding of the D protein, a putative resolvase of the F-factor of *Escherichia coli*. *Mol Gen Genet MGG* **237**: 206–214.
- Dominguez-Escobar, J., Chastanet, A., Crevenna, A.H., Fromion, V., Wedlich-Soldner, R., and Carballido-Lopez, R. (2011) Processive Movement of MreB-Associated Cell Wall Biosynthetic Complexes in Bacteria. *Science* **333**: 225–228.
- Donovan, C., Schwaiger, A., Krämer, R., and Bramkamp, M. (2010) Subcellular localization and characterization of the ParAB system from *Corynebacterium glutamicum*. *J Bacteriol* **192**: 3441–3451.
- Dubarry, N., Pasta, F., and Lane, D. (2006) ParABS systems of the four replicons of *Burkholderia cenocepacia*: new chromosome centromeres confer partition specificity. *J Bacteriol* **188**: 1489.
- Dunham, T.D., Xu, W., Funnell, B.E., and Schumacher, M.A. (2009) Structural basis for ADP-mediated transcriptional regulation by P1 and P7 ParA. *EMBO J* **28**: 1792–1802.
- Dupaigne, P., Tonthat, N.K., Espéli, O., Whitfill, T., Boccard, F., and Schumacher, M.A. (2012) Molecular basis for a protein-mediated DNA-bridging mechanism that functions in condensation of the *E. coli* chromosome. *Mol Cell* **48**: 560–571.
- Dworkin, J., and Losick, R. (2002) Does RNA polymerase help drive chromosome segregation in bacteria? *Proc Natl Acad Sci* **99**: 14089–14094.
- Ebersbach, G., Briegel, A., Jensen, G.J., and Jacobs-Wagner, C. (2008) A self-associating protein critical for chromosome attachment, division, and polar organization in *caulobacter*. *Cell* **134**: 956–968.
- Ebersbach, G., and Gerdes, K. (2001) The double par locus of virulence factor pB171: DNA segregation is correlated with oscillation of ParA. *Proc Natl Acad Sci* **98**: 15078.
- Ebersbach, G., and Gerdes, K. (2004) Bacterial mitosis: partitioning protein ParA oscillates in spiral-shaped structures and positions plasmids at mid-cell. *Mol Microbiol* **52**: 385–398.

-
- Ebersbach, G., Ringgaard, S., MU00F8ller-Jensen, J., Wang, Q., Sherratt, D.J., and Gerdes, K. (2006) Regular cellular distribution of plasmids by oscillating and filament-forming ParA ATPase of plasmid pB171. *Mol Microbiol* **61**: 1428–1442.
- Edgar, R., Chatteraj, D.K., and Yarmolinsky, M. (2001) Pairing of P1 plasmid partition sites by ParB. *Mol Microbiol* **42**: 1363–1370.
- Egan, E.S., and Waldor, M.K. (2003) Distinct replication requirements for the two *Vibrio cholerae* chromosomes. *Cell* **114**: 521–530.
- Eicken, C., Pennella, M.A., Chen, X., Koshlap, K.M., VanZile, M.L., Sacchettini, J.C., and Giedroc, D.P. (2003) A Metal–Ligand-mediated Intersubunit Allosteric Switch in Related SmtB/ArsR Zinc Sensor Proteins. *J Mol Biol* **333**: 683–695.
- Ent, F. van den, Møller-Jensen, J., Amos, L.A., Gerdes, K., and Löwe, J. (2002) F-actin-like filaments formed by plasmid segregation protein ParM. *EMBO J* **21**: 6935–6943.
- Erdmann, N., Petroff, T., and Funnell, B.E. (1999) Intracellular localization of P1 ParB protein depends on ParA and parS. *Proc Natl Acad Sci U S A* **96**: 14905–14910.
- Errington, J., Murray, H., and Wu, L.J. (2005) Diversity and redundancy in bacterial chromosome segregation mechanisms. *Philos Trans R Soc Lond B Biol Sci* **360**: 497–505.
- Fekete, R.A., and Chatteraj, D.K. (2005) A cis-acting sequence involved in chromosome segregation in *Escherichia coli*. *Mol Microbiol* **55**: 175–183.
- Figge, R.M., Easter, J., and Gober, J.W. (2003) Productive interaction between the chromosome partitioning proteins, ParA and ParB, is required for the progression of the cell cycle in *Caulobacter crescentus*. *Mol Microbiol* **47**: 1225–1237.
- Flärdh, K., and Buttner, M.J. (2009) *Streptomyces* morphogenetics: dissecting differentiation in a filamentous bacterium. *Nat Rev Microbiol* **7**: 36–49.
- Fogel, M.A., and Waldor, M.K. (2006) A dynamic, mitotic-like mechanism for bacterial chromosome segregation. *Genes Dev* **20**: 3269–3282.
- Friedman, S.A., and Austin, S.J. (1988) The P1 plasmid-partition system synthesizes two essential proteins from an autoregulated operon. *Plasmid* **19**: 103–112.
- Fung, E., Bouet, J.Y., and Funnell, B.E. (2001) Probing the ATP-binding site of P1 ParA: partition and repression have different requirements for ATP binding and hydrolysis. *EMBO J* **20**: 4901–4911.
- Funnell, B.E. (1988) Participation of *Escherichia coli* integration host factor in the P1 plasmid partition system. *Proc Natl Acad Sci* **85**: 6657.
- Funnell, B.E. (1988) Mini-P1 plasmid partitioning: excess ParB protein destabilizes plasmids containing the centromere parS. *J Bacteriol* **170**: 954–960.

- Funnell, B.E. (1991) The P1 plasmid partition complex at parS. The influence of Escherichia coli integration host factor and of substrate topology. *J Biol Chem* **266**: 14328–14337.
- Funnell, B.E., and Gagnier, L. (1993) The P1 plasmid partition complex at parS. II. Analysis of ParB protein binding activity and specificity. *J Biol Chem* **268**: 3616–3624.
- Gajiwala, K.S., and Burley, S.K. (2000) Winged helix proteins. *Curr Opin Struct Biol* **10**: 110–116.
- Galkin, V.E., Orlova, A., Rivera, C., Mullins, R.D., and Egelman, E.H. (2009) Structural Polymorphism of the ParM Filament and Dynamic Instability. *Structure* **17**: 1253–1264.
- Gallie, D.R., and Kado, C.I. (1987) Agrobacterium tumefaciens pTAR parA promoter region involved in autoregulation, incompatibility and plasmid partitioning. *J Mol Biol* **193**: 465–478.
- Garner, E.C., Bernard, R., Wang, W., Zhuang, X., Rudner, D.Z., and Mitchison, T. (2011) Coupled, Circumferential Motions of the Cell Wall Synthesis Machinery and MreB Filaments in B. subtilis. *Science* **333**: 222–225.
- Garner, E.C., Campbell, C.S., and Mullins, R.D. (2004) Dynamic instability in a DNA-segregating prokaryotic actin homolog. *Science* **306**: 1021–1025.
- Gelles, J., and Landick, R. (1998) RNA polymerase as a molecular motor. *Cell* **93**: 13–16.
- Gerdes, K. (2000) Toxin-antitoxin modules may regulate synthesis of macromolecules during nutritional stress. *J Bacteriol* **182**: 561–572.
- Gerdes, K., Gulyaev, A.P., Franch, T., Pedersen, K., and Mikkelsen, N.D. (1997) Antisense RNA-regulated programmed cell death. *Annu Rev Genet* **31**: 1–31.
- Gerdes, K., Howard, M., and Szardenings, F. (2010) Pushing and Pulling in Prokaryotic DNA Segregation. *Cell* **141**: 927–942.
- Gerdes, K., Larsen, J.E., and Molin, S. (1985) Stable inheritance of plasmid R1 requires two different loci. *J Bacteriol* **161**: 292.
- Gerdes, K., and Maisonneuve, E. (2012) Bacterial persistence and toxin-antitoxin Loci. *Annu Rev Microbiol* **66**: 103–123.
- Gerdes, K., MU00F8ller-Jensen, J., and Jensen, R.B. (2000) Plasmid and chromosome partitioning: surprises from phylogeny. *Mol Microbiol* **37**: 455–466.
- Gitai, Z., Dye, N.A., Reisenauer, A., Wachi, M., and Shapiro, L. (2005) MreB Actin-Mediated Segregation of a Specific Region of a Bacterial Chromosome. *Cell* **120**: 329–341.
- Glaser, P., Sharpe, M.E., Raether, B., Perego, M., Ohlsen, K., and Errington, J. (1997) Dynamic, mitotic-like behavior of a bacterial protein required for accurate chromosome partitioning. *Genes Dev* **11**: 1160–1168.

-
- Godfrin-Estevenon, A.-M., Pasta, F., and Lane, D. (2002) The parAB gene products of *Pseudomonas putida* exhibit partition activity in both *P. putida* and *Escherichia coli*. *Mol Microbiol* **43**: 39–49.
- Golovanov, A.P., Barillà, D., Golovanova, M., Hayes, F., and Lian, L.-Y. (2003) ParG, a protein required for active partition of bacterial plasmids, has a dimeric ribbon-helix-helix structure. *Mol Microbiol* **50**: 1141–1153.
- Gordon, G.S., Sitnikov, D., Webb, C.D., Teleman, A., Straight, A., Losick, R., *et al.* (1997) Chromosome and low copy plasmid segregation in *E. coli*: visual evidence for distinct mechanisms. *Cell* **90**: 1113–1121.
- Gordon, S., Rech, J., Lane, D., and Wright, A. (2004) Kinetics of plasmid segregation in *Escherichia coli*. *Mol Microbiol* **51**: 461–469.
- Grigoriev, P.S., and Lobočka, M.B. (2001) Determinants of segregational stability of the linear plasmid-prophage N15 of *Escherichia coli*. *Mol Microbiol* **42**: 355–368.
- Gruber, S., and Errington, J. (2009) Recruitment of condensin to replication origin regions by ParB/SpoOJ promotes chromosome segregation in *B. subtilis*. *Cell* **137**: 685–696.
- Hanai, R., Liu, R., Benedetti, P., Caron, P.R., Lynch, A.S., and Wang, J.C. (1996) Molecular dissection of a protein SopB essential for *Escherichia coli* F plasmid partition. *J Biol Chem* **271**: 17469–17475.
- Hao, J.-J., and Yarmolinsky, M. (2002) Effects of the P1 plasmid centromere on expression of P1 partition genes. *J Bacteriol* **184**: 4857–4867.
- Hatano, T., and Niki, H. (2010) Partitioning of P1 plasmids by gradual distribution of the ATPase ParA. *Mol Microbiol* **78**: 1182–1198.
- Hatano, T., Yamaichi, Y., and Niki, H. (2007) Oscillating focus of SopA associated with filamentous structure guides partitioning of F plasmid. *Mol Microbiol* **64**: 1198–1213.
- Havey, J.C., Vecchiarelli, A.G., and Funnell, B.E. (2012) ATP-regulated interactions between P1 ParA, ParB and non-specific DNA that are stabilized by the plasmid partition site, parS. *Nucleic Acids Res* **40**: 801–812.
- Hayakawa, Y., Murotsu, T., and Matsubara, K. (1985) Mini-F protein that binds to a unique region for partition of mini-F plasmid DNA. *J Bacteriol* **163**: 349–354.
- Hayes, F., and Austin, S. (1994) Topological scanning of the P1 plasmid partition site. *J Mol Biol* **243**: 190–198.
- Hayes, F., and Austin, S.J. (1993) Specificity determinants of the P1 and P7 plasmid centromere analogs. *Proc Natl Acad Sci U S A* **90**: 9228–9232.
- Hayes, F., and Barillà, D. (2006) The bacterial segrosome: a dynamic nucleoprotein machine for DNA trafficking and segregation. *Nat Rev Microbiol* **4**: 133–143.

- Hayes, F., Davis, M.A., and Austin, S.J. (1993) Fine-structure analysis of the P7 plasmid partition site. *J Bacteriol* **175**: 3443–3451.
- Hayes, F., Radnedge, L., Davis, M.A., and Austin, S.J. (1994) The homologous operons for P1 and P7 plasmid partition are autoregulated from dissimilar operator sites. *Mol Microbiol* **11**: 249–260.
- Hester, C.M., and Lutkenhaus, J. (2007) Soj (ParA) DNA binding is mediated by conserved arginines and is essential for plasmid segregation. *Proc Natl Acad Sci* **104**: 20326.
- Hiraga, S. (2000) Dynamic localization of bacterial and plasmid chromosomes. *Annu Rev Genet* **34**: 21–59.
- Hiraga, S., Niki, H., Ogura, T., Ichinose, C., Mori, H., Ezaki, B., and Jaffe, A. (1989) Chromosome partitioning in *Escherichia coli*: novel mutants producing anucleate cells. *J Bacteriol* **171**: 1496–1505.
- Hirano, M., Mori, H., Onogi, T., Yamazoe, M., Niki, H., Ogura, T., and Hiraga, S. (1998) Autoregulation of the partition genes of the mini-F plasmid and the intracellular localization of their products in *Escherichia coli*. *Mol Gen Genet MGG* **257**: 392–403.
- Hoch, J.A., and Mathews, J.L. (1973) Chromosomal location of pleiotropic negative sporulation mutations in *Bacillus subtilis*. *Genetics* **73**: 215–228.
- Hoischen, C., Bolshoy, A., Gerdes, K., and Diekmann, S. (2004) Centromere parC of plasmid R1 is curved. *Nucleic Acids Res* **32**: 5907–5915.
- Holmes, V.F., and Cozzarelli, N.R. (2000) Closing the ring: links between SMC proteins and chromosome partitioning, condensation, and supercoiling. *Proc Natl Acad Sci U S A* **97**: 1322–1324.
- Howard, M., and Gerdes, K. (2010) What is the mechanism of ParA-mediated DNA movement? *Mol Microbiol* **78**: 9–12.
- Hranueli, D., Piggot, P.J., and Mandelstam, J. (1974) Statistical estimate of the total number of operons specific for *Bacillus subtilis* sporulation. *J Bacteriol* **119**: 684–690.
- Huang, L., Yin, P., Zhu, X., Zhang, Y., and Ye, K. (2011) Crystal structure and centromere binding of the plasmid segregation protein ParB from pCXC100. *Nucleic Acids Res* **39**: 2954–2968.
- Hui, M.P., Galkin, V.E., Yu, X., Stasiak, A.Z., Stasiak, A., Waldor, M.K., and Egelman, E.H. (2010) ParA2, a *Vibrio cholerae* chromosome partitioning protein, forms left-handed helical filaments on DNA. *Proc Natl Acad Sci* **107**: 4590.
- Hunding, A., Ebersbach, G., and Gerdes, K. (2003) A Mechanism for ParB-dependent Waves of ParA, a Protein Related to DNA Segregation during Cell Division in Prokaryotes. *J Mol Biol* **329**: 35–43.

Hwang, L.C., Vecchiarelli, A.G., Han, Y.-W., Mizuuchi, M., Harada, Y., Funnell, B.E., and Mizuuchi, K. (2013) ParA-mediated plasmid partition driven by protein pattern self-organization. *EMBO J*.

Ireton, K., Gunther, N.W., 4th, and Grossman, A.D. (1994) spo0J is required for normal chromosome segregation as well as the initiation of sporulation in *Bacillus subtilis*. *J Bacteriol* **176**: 5320–5329.

JACOB, F., and BRENNER, S. (1963) [On the regulation of DNA synthesis in bacteria: the hypothesis of the replicon]. *Comptes Rendus Hebd Séances Académie Sci* **256**: 298–300.

JACOB, F., and WOLLMAN, E.L. (1956) [Processes of conjugation and recombination in *Escherichia coli*. I. Induction by conjugation or zygotic induction]. *Ann Inst Pasteur* **91**: 486–510.

Jagura-Burdzy, G., Macartney, D.P., Zatyka, M., Cunliffe, L., Cooke, D., Huggins, C., *et al.* (1999) Repression at a distance by the global regulator KorB of promiscuous IncP plasmids. *Mol Microbiol* **32**: 519–532.

Jakimowicz, D., Brzostek, A., Rumijowska-Galewicz, A., Zydek, P., Dolzblasz, A., Smulczyk-Krawczynszyn, A., *et al.* (2007) Characterization of the mycobacterial chromosome segregation protein ParB and identification of its target in *Mycobacterium smegmatis*. *Microbiology* **153**: 4050–4060.

Jakimowicz, D., Chater, K., and Zakrzewska-Czerwińska, J. (2002) The ParB protein of *Streptomyces coelicolor* A3(2) recognizes a cluster of parS sequences within the origin-proximal region of the linear chromosome. *Mol Microbiol* **45**: 1365–1377.

Jakimowicz, D., Gust, B., Zakrzewska-Czerwinska, J., and Chater, K.F. (2005) Developmental-Stage-Specific Assembly of ParB Complexes in *Streptomyces coelicolor* Hyphae. *J Bacteriol* **187**: 3572–3580.

Jakimowicz, D., and Wezel, G.P. van (2012) Cell division and DNA segregation in *Streptomyces*: how to build a septum in the middle of nowhere? *Mol Microbiol* **85**: 393–404.

Jensen, R.B., Dam, M., and Gerdes, K. (1994) Partitioning of plasmid R1. The parA operon is autoregulated by ParR and its transcription is highly stimulated by a downstream activating element. *J Mol Biol* **236**: 1299–1309.

Jensen, R.B., and Gerdes, K. (1999) Mechanism of DNA segregation in prokaryotes: ParM partitioning protein of plasmid R1 co-localizes with its replicon during the cell cycle. *EMBO J* **18**: 4076–4084.

Jensen, R.B., Lurz, R., and Gerdes, K. (1998) Mechanism of DNA segregation in prokaryotes: replicon pairing by parC of plasmid R1. *Proc Natl Acad Sci U S A* **95**: 8550–8555.

Jensen, R.B., and Shapiro, L. (1999) The *Caulobacter crescentus* smc gene is required for cell cycle progression and chromosome segregation. *Proc Natl Acad Sci* **96**: 10661.

- Jones, L.J., Carballido-López, R., and Errington, J. (2001) Control of cell shape in bacteria: helical, actin-like filaments in *Bacillus subtilis*. *Cell* **104**: 913–922.
- Jun, S., and Mulder, B. (2006) Entropy-driven spatial organization of highly confined polymers: lessons for the bacterial chromosome. *Proc Natl Acad Sci U S A* **103**: 12388–12393.
- Jun, S., and Wright, A. (2010) Entropy as the driver of chromosome segregation. *Nat Rev Microbiol* **8**: 600–607.
- Junier, I., Boccard, F., and Espeli, O. (2013) Polymer modeling of the *E. coli* genome reveals the involvement of locus positioning and macrodomain structuring for the control of chromosome conformation and segregation. *Nucleic Acids Res* **42**: 1461–1473.
- Kaimer, C., and Graumann, P.L. (2011) Players between the worlds: multifunctional DNA translocases. *Curr Opin Microbiol* **14**: 719–725.
- Kaimer, C., Schenk, K., and Graumann, P.L. (2011) Two DNA Translocases Synergistically Affect Chromosome Dimer Resolution in *Bacillus subtilis*. *J Bacteriol* **193**: 1334–1340.
- Kallioma-Sanford, A.K., Rodriguez-Castañeda, F.A., McLeod, B.N., Latorre-Roselló, V., Smith, J.H., Reimann, J., *et al.* (2012) Chromosome segregation in Archaea mediated by a hybrid DNA partition machine. *Proc Natl Acad Sci*
<http://www.pnas.org/content/early/2012/02/14/1113384109>. Accessed February 27, 2012.
- Kalnin, K., Stegalkina, S., and Yarmolinsky, M. (2000) pTAR-encoded proteins in plasmid partitioning. *J Bacteriol* **182**: 1889–1894.
- Khan, S.A. (2005) Plasmid rolling-circle replication: highlights of two decades of research. *Plasmid* **53**: 126–136.
- Khare, D., Ziegelin, G., Lanka, E., and Heinemann, U. (2004) Sequence-specific DNA binding determined by contacts outside the helix-turn-helix motif of the ParB homolog KorB. *Nat Struct Mol Biol* **11**: 656–663.
- Kim, H.J., Calcutt, M.J., Schmidt, F.J., and Chater, K.F. (2000) Partitioning of the linear chromosome during sporulation of *Streptomyces coelicolor* A3 (2) involves an oriC-linked parAB locus. *J Bacteriol* **182**: 1313.
- Kim, S.K., and Wang, J.C. (1998) Localization of F plasmid SopB protein to positions near the poles of *Escherichia coli* cells. *Proc Natl Acad Sci U S A* **95**: 1523–1527.
- Kim, S.-K., and Wang, J.C. (1999) Gene silencing via protein-mediated subcellular localization of DNA. *Proc Natl Acad Sci* **96**: 8557–8561.
- Koonin, E.V. (1993) A superfamily of ATPases with diverse functions containing either classical or deviant ATP-binding motif. *J Mol Biol* **229**: 1165–1174.
- Kruse, T., Blagoev, B., Løbner-Olesen, A., Wachi, M., Sasaki, K., Iwai, N., *et al.* (2006) Actin homolog MreB and RNA polymerase interact and are both required for chromosome segregation in *Escherichia coli*. *Genes Dev* **20**: 113–124.

Kusiak, M., Gapczynska, A., Plochocka, D., Thomas, C.M., and Jagura-Burdzy, G. (2011) Binding and Spreading of ParB on DNA Determine Its Biological Function in *Pseudomonas aeruginosa*. *J Bacteriol* **193**: 3342–3355.

Kusukawa, N., Mori, H., Kondo, A., and Hiraga, S. (1987) Partitioning of the F plasmid: overproduction of an essential protein for partition inhibits plasmid maintenance. *Mol Gen Genet MGG* **208**: 365–372.

Kwong, S.M., Yeo, C.C., and Poh, C.L. (2001) Molecular analysis of the pRA2 partitioning region: ParB autoregulates parAB transcription and forms a nucleoprotein complex with the plasmid partition site, parS. *Mol Microbiol* **40**: 621–633.

Lane, D., Feyter, R. de, Kennedy, M., Phua, S.H., and Semon, D. (1986) D protein of miniF plasmid acts as a repressor of transcription and as a site-specific resolvase. *Nucleic Acids Res* **14**: 9713–9728.

Lane, D., Rothenbuehler, R., Merrilat, A.M., and Aiken, C. (1987) Analysis of the F plasmid centromere. *Mol Gen Genet MGG* **207**: 406–412.

Larsen, R.A., Cusumano, C., Fujioka, A., Lim-Fong, G., Patterson, P., and Pogliano, J. (2007) Treadmilling of a prokaryotic tubulin-like protein, TubZ, required for plasmid stability in *Bacillus thuringiensis*. *Genes Dev* **21**: 1340–1352.

Lasocki, K., Bartosik, A.A., Mierzejewska, J., Thomas, C.M., and Jagura-Burdzy, G. (2007) Deletion of the parA (soj) homologue in *Pseudomonas aeruginosa* causes ParB instability and affects growth rate, chromosome segregation, and motility. *J Bacteriol* **189**: 5762.

LEDERBERG, J. (1952) Cell genetics and hereditary symbiosis. *Physiol Rev* **32**: 403–430.

Lee, M.-J., Liu, C.-H., Wang, S.-Y., Huang, C.-T., and Huang, H. (2006) Characterization of the Soj/Spo0J chromosome segregation proteins and identification of putative parS sequences in *Helicobacter pylori*. *Biochem Biophys Res Commun* **342**: 744–750.

Lee, P.S., and Grossman, A.D. (2006) The chromosome partitioning proteins Soj (ParA) and Spo0J (ParB) contribute to accurate chromosome partitioning, separation of replicated sister origins, and regulation of replication initiation in *Bacillus subtilis*. *Mol Microbiol* **60**: 853–869.

Lemon, K.P., and Grossman, A.D. (2001) The extrusion-capture model for chromosome partitioning in bacteria. *Genes Dev* **15**: 2031–2041.

Lemonnier, M., Bouet, J.Y., Libante, V., and Lane, D. (2000) Disruption of the F plasmid partition complex in vivo by partition protein SopA. *Mol Microbiol* **38**: 493–505.

Leonard, T.A., Butler, P.J., and Löwe, J. (2005) Bacterial chromosome segregation: structure and DNA binding of the Soj dimer--a conserved biological switch. *EMBO J* **24**: 270–282.

Leonard, T.A., Butler, P.J.G., and Löwe, J. (2004) Structural analysis of the chromosome segregation protein Spo0J from *Thermus thermophilus*. *Mol Microbiol* **53**: 419–432.

- Lesterlin, C., Barre, F.-X., and Cornet, F. (2004) Genetic recombination and the cell cycle: what we have learned from chromosome dimers. *Mol Microbiol* **54**: 1151–1160.
- Li, Y., and Austin, S. (2002) The P1 plasmid in action: time-lapse photomicroscopy reveals some unexpected aspects of plasmid partition. *Plasmid* **48**: 174–178.
- Libante, V., Thion, L., and Lane, D. (2001) Role of the ATP-binding site of SopA protein in partition of the F plasmid. *J Mol Biol* **314**: 387–399.
- Lim, G.E., Derman, A.I., and Pogliano, J. (2005) Bacterial DNA segregation by dynamic SopA polymers. *Proc Natl Acad Sci U S A* **102**: 17658–17663.
- Lin, D.C., and Grossman, A.D. (1998) Identification and characterization of a bacterial chromosome partitioning site. *Cell* **92**: 675–685.
- Lin, D.C., Levin, P.A., and Grossman, A.D. (1997) Bipolar localization of a chromosome partition protein in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* **94**: 4721–4726.
- Livny, J., Yamaichi, Y., and Waldor, M.K. (2007) Distribution of centromere-like parS sites in bacteria: insights from comparative genomics. *J Bacteriol* **189**: 8693–8703.
- Lobocka, M., and Yarmolinsky, M. (1996) P1 plasmid partition: a mutational analysis of ParB. *J Mol Biol* **259**: 366–382.
- Lobry, J.R. (1996) Origin of replication of *Mycoplasma genitalium*. *Science* **272**: 745–746.
- Ludtke, D.N., Eichorn, B.G., and Austin, S.J. (1989) Plasmid-partition functions of the P7 prophage. *J Mol Biol* **209**: 393–406.
- Lukaszewicz, M., Kostelidou, K., Bartosik, A.A., Cooke, G.D., Thomas, C.M., and Jagura-Burdzy, G. (2002) Functional dissection of the ParB homologue (KorB) from IncP-1 plasmid RK2. *Nucleic Acids Res* **30**: 1046–1055.
- Lynch, A.S., and Wang, J.C. (1994) Use of an inducible site-specific recombinase to probe the structure of protein-DNA complexes involved in F plasmid partition in *Escherichia coli*. *J Mol Biol* **236**: 679–684.
- Lynch, A.S., and Wang, J.C. (1995) SopB protein-mediated silencing of genes linked to the sopC locus of *Escherichia coli* F plasmid. *Proc Natl Acad Sci* **92**: 1896.
- Maloney, E., Madiraju, M., and Rajagopalan, M. (2009) Overproduction and localization of *Mycobacterium tuberculosis* ParA and ParB proteins. *Tuberc Edinb Scotl* **89 Suppl 1**: S65–69.
- Marko, J.F. (2009) Linking topology of tethered polymer rings with applications to chromosome segregation and estimation of the knotting length. *Phys Rev E Stat Nonlin Soft Matter Phys* **79**: 051905.
- Marston, A.L., and Errington, J. (1999) Dynamic Movement of the ParA-like Soj Protein of *B. subtilis* and Its Dual Role in Nucleoid Organization and Developmental Regulation. *Mol Cell* **4**: 673–682.

-
- Martin, K.A., Friedman, S.A., and Austin, S.J. (1987) Partition site of the P1 plasmid. *Proc Natl Acad Sci* **84**: 8544.
- McEachern, M.J., Bott, M.A., Tooker, P.A., and Helinski, D.R. (1989) Negative control of plasmid R6K replication: possible role of intermolecular coupling of replication origins. *Proc Natl Acad Sci U S A* **86**: 7942–7946.
- Mercier, R., Petit, M.-A., Schbath, S., Robin, S., El Karoui, M., Boccard, F., and Espéli, O. (2008) The MatP/matS Site-Specific System Organizes the Terminus Region of the E. coli Chromosome into a Macrodomain. *Cell* **135**: 475–485.
- Messer, W. (2002) The bacterial replication initiator DnaA. DnaA and oriC, the bacterial mode to initiate DNA replication. *FEMS Microbiol Rev* **26**: 355–374.
- Minnen, A., Attaiach, L., Thon, M., Gruber, S., and Veening, J.-W. (2011) SMC is recruited to oriC by ParB and promotes chromosome segregation in *Streptococcus pneumoniae*. *Mol Microbiol* **81**: 676–688.
- Mohl, D.A., Easter Jr, J., and Gober, J.W. (2001) The chromosome partitioning protein, ParB, is required for cytokinesis in *Caulobacter crescentus*. *Mol Microbiol* **42**: 741–755.
- Mohl, D.A., and Gober, J.W. (1997) Cell cycle-dependent polar localization of chromosome partitioning proteins in *Caulobacter crescentus*. *Cell* **88**: 675–684.
- Møller-Jensen, J., Borch, J., Dam, M., Jensen, R.B., Roepstorff, P., and Gerdes, K. (2003) Bacterial Mitosis:: ParM of Plasmid R1 Moves Plasmid DNA by an Actin-like Insertional Polymerization Mechanism. *Mol Cell* **12**: 1477–1487.
- Møller-Jensen, J., Jensen, R.B., Löwe, J., and Gerdes, K. (2002) Prokaryotic DNA segregation by an actin-like filament. *EMBO J* **21**: 3119–3127.
- Møller-Jensen, J., Ringgaard, S., Mercogliano, C.P., Gerdes, K., and Löwe, J. (2007) Structural analysis of the ParR/parC plasmid partition complex. *EMBO J* **26**: 4413–4422.
- Mori, H., Kondo, A., Ohshima, A., Ogura, T., and Hiraga, S. (1986) Structure and function of the F plasmid genes essential for partitioning. *J Mol Biol* **192**: 1–15.
- Mori, H., Mori, Y., Ichinose, C., Niki, H., Ogura, T., Kato, A., and Hiraga, S. (1989) Purification and characterization of SopA and SopB proteins essential for F plasmid partitioning. *J Biol Chem* **264**: 15535–15541.
- Morrison, P.F., and Chatteraj, D.K. (2004) Replication of a unit-copy plasmid F in the bacterial cell cycle: a replication rate function analysis. *Plasmid* **52**: 13–30.
- Murayama, K., Orth, P., Hoz, A.B. de la, Alonso, J.C., and Saenger, W. (2001) Crystal structure of ω transcriptional repressor encoded by *Streptococcus pyogenes* plasmid pSM19035 at 1.5 Å resolution. *J Mol Biol* **314**: 789–796.
- Murray, H., and Errington, J. (2008) Dynamic control of the DNA replication initiation protein DnaA by Soj/ParA. *Cell* **135**: 74–84.

- Murray, H., Ferreira, H., and Errington, J. (2006) The bacterial chromosome segregation protein Spo0J spreads along DNA from parS nucleation sites. *Mol Microbiol* **61**: 1352–1361.
- Nadanaciva, S., Weber, J., Wilke-Mounts, S., and Senior, A.E. (1999) Importance of F₁-ATPase Residue α -Arg-376 for Catalytic Transition State Stabilization[†]. *Biochemistry (Mosc)* **38**: 15493–15499.
- Nasmyth, K., and Haering, C.H. (2005) THE STRUCTURE AND FUNCTION OF SMC AND KLEISIN COMPLEXES. *Annu Rev Biochem* **74**: 595–648.
- Ni, L., Xu, W., Kumaraswami, M., and Schumacher, M.A. (2010) Plasmid protein TubR uses a distinct mode of HTH-DNA binding and recruits the prokaryotic tubulin homolog TubZ to effect DNA partition. *Proc Natl Acad Sci* **107**: 11763–11768.
- Nielsen, H.J., Li, Y., Youngren, B., Hansen, F.G., and Austin, S. (2006) Progressive segregation of the Escherichia coli chromosome. *Mol Microbiol* **61**: 383–393.
- Nielsen, H.J., Youngren, B., Hansen, F.G., and Austin, S. (2007) Dynamics of Escherichia coli Chromosome Segregation during Multifork Replication. *J Bacteriol* **189**: 8660–8666.
- Niki, H., and Hiraga, S. (1997) Subcellular distribution of actively partitioning F plasmid during the cell division cycle in E. coli. *Cell* **90**: 951–957.
- Niki, H., and Hiraga, S. (1998) Polar localization of the replication origin and terminus in Escherichia coli nucleoids during chromosome partitioning. *Genes Dev* **12**: 1036.
- Niki, H., Jaffé, A., Imamura, R., Ogura, T., and Hiraga, S. (1991) The new gene mukB codes for a 177 kd protein with coiled-coil domains involved in chromosome partitioning of E. coli. *EMBO J* **10**: 183.
- Niki, H., Yamaichi, Y., and Hiraga, S. (2000) Dynamic organization of chromosomal DNA in Escherichia coli. *Genes Dev* **14**: 212–223.
- Nordström, K., and Wagner, E.G. (1994) Kinetic aspects of control of plasmid replication by antisense RNA. *Trends Biochem Sci* **19**: 294–300.
- Ogura, T., and Hiraga, S. (1983) Partition mechanism of F plasmid: two plasmid gene-encoded products and a cis-acting region are involved in partition. *Cell* **32**: 351–360.
- Ogura, T., Miki, T., and Hiraga, S. (1980) Copy-number mutants of the plasmid carrying the replication origin of the Escherichia coli chromosome: evidence for a control region of replication. *Proc Natl Acad Sci U S A* **77**: 3993–3997.
- Oliva, M.A., Cordell, S.C., and Löwe, J. (2004) Structural insights into FtsZ protofilament formation. *Nat Struct Mol Biol* **11**: 1243–1250.
- Orlova, A., Garner, E.C., Galkin, V.E., Heuser, J., Mullins, R.D., and Egelman, E.H. (2007) The structure of bacterial ParM filaments. *Nat Struct Mol Biol* **14**: 921–926.

-
- Painbeni, E., Caroff, M., and Rouviere-Yaniv, J. (1997) Alterations of the outer membrane composition in *Escherichia coli* lacking the histone-like protein HU. *Proc Natl Acad Sci U S A* **94**: 6712.
- Passot, F.M., Calderon, V., Fichant, G., Lane, D., and Pasta, F. (2012) Centromere binding and evolution of chromosomal partition systems in the Burkholderiales. *J Bacteriol* **194**: 3426–3436.
- Pennella, M.A., and Giedroc, D.P. (2005) Structural Determinants of Metal Selectivity in Prokaryotic Metal-responsive Transcriptional Regulators. *Biometals* **18**: 413–428.
- Piggot, P.J., and Coote, J.G. (1976) Genetic aspects of bacterial endospore formation. *Bacteriol Rev* **40**: 908–962.
- Pillet, F., Sanchez, A., Lane, D., Anton Leberre, V., and Bouet, J.-Y. (2011) Centromere binding specificity in assembly of the F plasmid partition complex. *Nucleic Acids Res* <http://nar.oxfordjournals.org/content/early/2011/06/07/nar.gkr457.abstract>. Accessed July 21, 2011.
- Pogliano, J., Ho, T.Q., Zhong, Z., and Helinski, D.R. (2001) Multicopy plasmids are clustered and localized in *Escherichia coli*. *Proc Natl Acad Sci U S A* **98**: 4486–4491.
- Polka, J.K., Kollman, J.M., Agard, D.A., and Mullins, R.D. (2009) The Structure and Assembly Dynamics of Plasmid Actin AlfA Imply a Novel Mechanism of DNA Segregation. *J Bacteriol* **191**: 6219–6230.
- Popp, D., Narita, A., Ghoshdastider, U., Maeda, K., Maéda, Y., Oda, T., *et al.* (2010) Polymeric structures and dynamic properties of the bacterial actin AlfA. *J Mol Biol* **397**: 1031–1041.
- Popp, D., Narita, A., Oda, T., Fujisawa, T., Matsuo, H., Nitanaï, Y., *et al.* (2008) Molecular structure of the ParM polymer and the mechanism leading to its nucleotide-driven dynamic instability. *EMBO J* **27**: 570–579.
- Popp, D., Xu, W., Narita, A., Brzoska, A.J., Skurray, R.A., Firth, N., *et al.* (2010) Structure and Filament Dynamics of the pSK41 Actin-like ParM Protein: IMPLICATIONS FOR PLASMID DNA SEGREGATION. *J Biol Chem* **285**: 10130–10140.
- Postow, L. (2004) Topological domain structure of the *Escherichia coli* chromosome. *Genes Dev* **18**: 1766–1779.
- Pratto, F., Cicek, A., Weihofen, W.A., Lurz, R., Saenger, W., and Alonso, J.C. (2008) *Streptococcus pyogenes* pSM19035 requires dynamic assembly of ATP-bound ParA and ParB on parS DNA during plasmid segregation. *Nucleic Acids Res* **36**: 3676–3689.
- Pratto, F., Suzuki, Y., Takeyasu, K., and Alonso, J.C. (2009) Single-molecule analysis of proteinxDNA complexes formed during partition of newly replicated plasmid molecules in *Streptococcus pyogenes*. *J Biol Chem* **284**: 30298–30306.
- Ptacin, J.L., Lee, S.F., Garner, E.C., Toro, E., Eckart, M., Comolli, L.R., *et al.* (2010) A spindle-like apparatus guides bacterial chromosome segregation. *Nat Cell Biol* **12**: 791–798.

- Ptacin, J.L., Nollmann, M., Becker, E.C., Cozzarelli, N.R., Pogliano, K., and Bustamante, C. (2008) Sequence-directed DNA export guides chromosome translocation during sporulation in *Bacillus subtilis*. *Nat Struct Mol Biol* **15**: 485–493.
- Quisel, J.D., Lin, D.C.-H., and Grossman, A.D. (1999) Control of Development by Altered Localization of a Transcription Factor in *B. subtilis*. *Mol Cell* **4**: 665–672.
- Radnedge, L., Youngren, B., Davis, M., and Austin, S. (1998) Probing the structure of complex macromolecular interactions by homolog specificity scanning: the P1 and P7 plasmid partition systems. *EMBO J* **17**: 6076–6085.
- Ravin, N.V., Rech, J., and Lane, D. (2003) Mapping of functional domains in F plasmid partition proteins reveals a bipartite SopB-recognition domain in SopA. *J Mol Biol* **329**: 875–889.
- Reyes-Lamothe, R., Nicolas, E., and Sherratt, D.J. (2012) Chromosome Replication and Segregation in Bacteria. *Annu Rev Genet* <http://www.ncbi.nlm.nih.gov/pubmed/22934648>. Accessed September 3, 2012.
- Reyes-Lamothe, R., Possoz, C., Danilova, O., and Sherratt, D.J. (2008) Independent Positioning and Action of *Escherichia coli* Replisomes in Live Cells. *Cell* **133**: 90–102.
- Rice, P.A., Yang, S., Mizuuchi, K., and Nash, H.A. (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell* **87**: 1295–1306.
- Ringgaard, S., Ebersbach, G., Borch, J., and Gerdes, K. (2006) Regulatory Cross-talk in the Double par Locus of Plasmid pB171. *J Biol Chem* **282**: 3134–3145.
- Ringgaard, S., Löwe, J., and Gerdes, K. (2007) Centromere pairing by a plasmid-encoded type I ParB protein. *J Biol Chem* **282**: 28216–28225.
- Ringgaard, S., Zon, J. van, Howard, M., and Gerdes, K. (2009) Movement and equipositioning of plasmids by ParA filament disassembly. *Proc Natl Acad Sci U S A* **106**: 19369–19374.
- Rittinger, K., Walker, P.A., Eccleston, J.F., Smerdon, S.J., and Gamblin, S.J. (1997) Structure at 1.65 Å of RhoA and its GTPase-activating protein in complex with a transition-state analogue. *Nature* **389**: 758–762.
- Rocha, E.P.C., Touchon, M., and Feil, E.J. (2006) Similar compositional biases are caused by very different mutational effects. *Genome Res* **16**: 1537–1547.
- Rodionov, O., Lobocka, M., and Yarmolinsky, M. (1999) Silencing of genes flanking the P1 plasmid centromere. *Science* **283**: 546–549.
- Rodionov, O., and Yarmolinsky, M. (2004) Plasmid partitioning and the spreading of P1 partition protein ParB. *Mol Microbiol* **52**: 1215–1223.
- Saint-Dic, D., Frushour, B.P., Kehrl, J.H., and Kahng, L.S. (2006) A parA homolog selectively influences positioning of the large chromosome origin in *Vibrio cholerae*. *J Bacteriol* **188**: 5626–5631.

Sakai, H., and Komano, T. (1996) DNA replication of IncQ broad-host-range plasmids in gram-negative bacteria. *Biosci Biotechnol Biochem* **60**: 377–382.

Salanoubat, M., Genin, S., Artiguenave, F., Gouzy, J., Mangenot, S., Arlat, M., *et al.* (2002) Genome sequence of the plant pathogen *Ralstonia solanacearum*. *Nature* **415**: 497–502.

Salje, J. (2010) Plasmid segregation: how to survive as an extra piece of DNA. *Crit Rev Biochem Mol Biol* **45**: 296–317.

Salje, J., and Löwe, J. (2008) Bacterial actin: architecture of the ParMRC plasmid DNA partitioning complex. *EMBO J* **27**: 2230–2238.

Sanchez, A., Rech, J., Gasc, C., and Bouet, J.-Y. (2013) Insight into centromere-binding properties of ParB proteins: a secondary binding motif is essential for bacterial genome maintenance. *Nucleic Acids Res* **41**: 3094–3103.

Scheffzek, K. (1997) The Ras-RasGAP Complex: Structural Basis for GTPase Activation and Its Loss in Oncogenic Ras Mutants. *Science* **277**: 333–338.

Scheffzek, K., Ahmadian, M.R., and Wittinghofer, A. (1998) GTPase-activating proteins: helping hands to complement an active site. *Trends Biochem Sci* **23**: 257–262.

Scherzinger, E., Haring, V., Lurz, R., and Otto, S. (1991) Plasmid RSF1010 DNA replication in vitro promoted by purified RSF1010 RepA, RepB and RepC proteins. *Nucleic Acids Res* **19**: 1203–1211.

Schumacher, M. (2008) Structural biology of plasmid partition: uncovering the molecular mechanisms of DNA segregation. *Biochem J* **412**: 1–18.

Schumacher, M.A. (2007) Structural biology of plasmid segregation proteins. *Curr Opin Struct Biol* **17**: 103–109.

Schumacher, M.A., and Funnell, B.E. (2005) Structures of ParB bound to DNA reveal mechanism of partition complex formation. *Nature* **438**: 516–519.

Schumacher, M.A., Glover, T.C., Brzoska, A.J., Jensen, S.O., Dunham, T.D., Skurray, R.A., and Firth, N. (2007) Segrosome structure revealed by a complex of ParR with centromere DNA. *Nature* **450**: 1268–1271.

Schumacher, M.A., Mansoor, A., and Funnell, B.E. (2007) Structure of a four-way bridged ParB-DNA complex provides insight into P1 segrosome assembly. *J Biol Chem* **282**: 10456–10464.

Schumacher, M.A., Piro, K.M., and Xu, W. (2010) Insight into F plasmid DNA segregation revealed by structures of SopB and SopB-DNA complexes. *Nucleic Acids Res* **38**: 4514–4526.

Schwartz, M.A., and Shapiro, L. (2011) An SMC ATPase mutant disrupts chromosome segregation in *Caulobacter*: Role of the SMC ATPase in chromosome segregation. *Mol Microbiol* **82**: 1359–1374.

- Ah-Seng, Y., Lopez, F., Pasta, F., Lane, D., and Bouet, J.-Y. (2009) Dual Role of DNA in Regulating ATP Hydrolysis by the SopA Partition Protein. *J Biol Chem* **284**: 30067–30075.
- Ah-Seng, Y., Rech, J., Lane, D., and Bouet, J.-Y. (2013) Defining the Role of ATP Hydrolysis in Mitotic Segregation of Bacterial Plasmids. *PLoS Genet* **9**: e1003956.
- Sengupta, M., Nielsen, H.J., Youngren, B., and Austin, S. (2009) P1 Plasmid Segregation: Accurate Redistribution by Dynamic Plasmid Pairing and Separation. *J Bacteriol* **192**: 1175–1183.
- Sergueev, K., Dabrazhynetskaya, A., and Austin, S. (2005) Plasmid partition system of the P1par family from the pWR100 virulence plasmid of *Shigella flexneri*. *J Bacteriol* **187**: 3369–3373.
- Sharpe, M.E., and Errington, J. (1996) The *Bacillus subtilis* *soj-spo0J* locus is required for a centromere-like function involved in prespore chromosome partitioning. *Mol Microbiol* **21**: 501–509.
- Shebelut, C.W., Guberman, J.M., Teeffelen, S. van, Yakhnina, A.A., and Gitai, Z. (2010) *Caulobacter* chromosome segregation is an ordered multistep process. *Proc Natl Acad Sci U S A* **107**: 14194–14198.
- Slater, F.R., Bailey, M.J., Tett, A.J., and Turner, S.L. (2008) Progress towards understanding the fate of plasmids in bacterial communities. *FEMS Microbiol Ecol* **66**: 3–13.
- Solar, G. Del, and Espinosa, M. (2000) Plasmid copy number control: an ever-growing story. *Mol Microbiol* **37**: 492–500.
- Solar, G. Del, Giraldo, R., Ruiz-Echevarria, M.J., Espinosa, M., and Diaz-Orejas, R. (1998) Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* **62**: 434.
- Srivastava, P., and Chattoraj, D.K. (2007) Selective chromosome amplification in *Vibrio cholerae*. *Mol Microbiol* **66**: 1016–1028.
- Steiner, W.W., and Kuempel, P.L. (1998) Sister Chromatid Exchange Frequencies in *Escherichia coli* Analyzed by Recombination at the *dif* Resolvase Site. *J Bacteriol* **180**: 6269–6275.
- Stouf, M., Meile, J.-C., and Cornet, F. (2013) FtsK actively segregates sister chromosomes in *Escherichia coli*. *Proc Natl Acad Sci* <http://www.pnas.org/cgi/doi/10.1073/pnas.1304080110>. Accessed June 24, 2013.
- Sullivan, N.L., Marquis, K.A., and Rudner, D.Z. (2009) Recruitment of SMC by ParB-parS Organizes the Origin Region and Promotes Efficient Chromosome Segregation. *Cell* **137**: 697–707.
- Summers, D.K., and Sherratt, D.J. (1984) Multimerization of high copy number plasmids causes instability: ColE1 encodes a determinant essential for plasmid monomerization and stability. *Cell* **36**: 1097–1103.

-
- Surtees, J.A. (2001) The DNA Binding Domains of P1 ParB and the Architecture of the P1 Plasmid Partition Complex. *J Biol Chem* **276**: 12385–12394.
- Surtees, J.A., and Funnell, B.E. (1999) P1 ParB domain structure includes two independent multimerization domains. *J Bacteriol* **181**: 5898.
- Szardenings, F., Guymer, D., and Gerdes, K. (2011) ParA ATPases can move and position DNA and subcellular structures. *Curr Opin Microbiol* **14**: 712–718.
- Tang, M., Bideshi, D.K., Park, H.-W., and Federici, B.A. (2006) Minireplicon from pBtoxis of *Bacillus thuringiensis* subsp. *israelensis*. *Appl Environ Microbiol* **72**: 6948–6954.
- Tang, M., Bideshi, D.K., Park, H.-W., and Federici, B.A. (2007) Iteron-Binding ORF157 and FtsZ-Like ORF156 Proteins Encoded by pBtoxis Play a Role in Its Replication in *Bacillus thuringiensis* subsp. *israelensis*. *J Bacteriol* **189**: 8053–8058.
- Thanbichler, M., and Shapiro, L. (2008) Getting organized — how bacterial cells move proteins and DNA. *Nat Rev Microbiol* **6**: 28–40.
- Timmis, K., Cabello, F., and Cohen, S.N. (1975) Cloning, isolation, and characterization of replication regions of complex plasmid genomes. *Proc Natl Acad Sci U S A* **72**: 2242–2246.
- Toro, E., and Shapiro, L. (2010) Bacterial chromosome organization and segregation. *Cold Spring Harb Perspect Biol* **2**: a000349.
- Typas, A., Banzhaf, M., Gross, C.A., and Vollmer, W. (2011) From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat Rev Microbiol* <http://www.nature.com/doifinder/10.1038/nrmicro2677>. Accessed February 18, 2014.
- Valens, M., Penaud, S., Rossignol, M., Cornet, F., and Boccard, F. (2004) Macrodomain organization of the *Escherichia coli* chromosome. *EMBO J* **23**: 4330–4341.
- Vecchiarelli, A.G., Han, Y.-W., Tan, X., Mizuuchi, M., Ghirlando, R., Biertümpfel, C., *et al.* (2010) ATP control of dynamic P1 ParA-DNA interactions: a key role for the nucleoid in plasmid partition. *Mol Microbiol* **78**: 78–91.
- Vecchiarelli, A.G., Havey, J.C., Ing, L.L., Wong, E.O.Y., Waples, W.G., and Funnell, B.E. (2013) Dissection of the ATPase Active Site of P1 ParA Reveals Multiple Active Forms Essential for Plasmid Partition. *J Biol Chem* **288**: 17823–17831.
- Vecchiarelli, A.G., Hwang, L.C., and Mizuuchi, K. (2013) Cell-free study of F plasmid partition provides evidence for cargo transport by a diffusion-ratchet mechanism. *Proc Natl Acad Sci U S A* .
- Vecchiarelli, A.G., Mizuuchi, K., and Funnell, B.E. (2012) Surfing biological surfaces: exploiting the nucleoid for partition and transport in bacteria. *Mol Microbiol* n/a–n/a.
- Vecchiarelli, A.G., Neuman, K.C., and Mizuuchi, K. (2014) A propagating ATPase gradient drives transport of surface-confined cellular cargo. *Proc Natl Acad Sci U S A* .

- Viollier, P.H., Thanbichler, M., McGrath, P.T., West, L., Meewan, M., McAdams, H.H., and Shapiro, L. (2004) Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. *Proc Natl Acad Sci U S A* **101**: 9257.
- Volkov, A., Mascarenhas, J., Andrei-Selmer, C., Ulrich, H.D., and Graumann, P.L. (2003) A Prokaryotic Condensin/Cohesin-Like Complex Can Actively Compact Chromosomes from a Single Position on the Nucleoid and Binds to DNA as a Ring-Like Structure. *Mol Cell Biol* **23**: 5638–5650.
- Wagner, E.G., and Simons, R.W. (1994) Antisense RNA control in bacteria, phages, and plasmids. *Annu Rev Microbiol* **48**: 713–742.
- Wang, J.C. (2002) Cellular roles of dna topoisomerases: a molecular perspective. *Nat Rev Mol Cell Biol* **3**: 430–440.
- Wang, L., Yu, Y., He, X., Zhou, X., Deng, Z., Chater, K.F., and Tao, M. (2007) Role of an FtsK-Like Protein in Genetic Stability in *Streptomyces coelicolor* A3(2). *J Bacteriol* **189**: 2310–2318.
- Wang, M.D. (1998) Force and Velocity Measured for Single Molecules of RNA Polymerase. *Science* **282**: 902–907.
- Wang, W., Li, G.-W., Chen, C., Xie, X.S., and Zhuang, X. (2011) Chromosome organization by a nucleoid-associated protein in live bacteria. *Science* **333**: 1445–1449.
- Wang, X., Possoz, C., and Sherratt, D.J. (2005) Dancing around the divisome: asymmetric chromosome segregation in *Escherichia coli*. *Genes Dev* **19**: 2367–2377.
- Wang, X., Reyes-Lamothe, R., and Sherratt, D.J. (2008) Visualizing genetic loci and molecular machines in living bacteria. *Biochem Soc Trans* **36**: 749–753.
- Watanabe, E., Wachi, M., Yamasaki, M., and Nagai, K. (1992) ATPase activity of SopA, a protein essential for active partitioning of F plasmid. *Mol Gen Genet MGG* **234**: 346–352.
- Webb, C.D., Graumann, P.L., Kahana, J.A., Teleman, A.A., Silver, P.A., and Losick, R. (1998) Use of time-lapse microscopy to visualize rapid movement of the replication origin region of the chromosome during the cell cycle in *Bacillus subtilis*. *Mol Microbiol* **28**: 883–892.
- Webb, C.D., Teleman, A., Gordon, S., Straight, A., Belmont, A., Lin, D.C., *et al.* (1997) Bipolar localization of the replication origin regions of chromosomes in vegetative and sporulating cells of *B. subtilis*. *Cell* **88**: 667–674.
- Weihofen, W.A., Cicek, A., Pratto, F., Alonso, J.C., and Saenger, W. (2006) Structures of omega repressors bound to direct and inverted DNA repeats explain modulation of transcription. *Nucleic Acids Res* **34**: 1450–1458.
- White, C.L., and Gober, J.W. (2012) MreB: pilot or passenger of cell wall synthesis? *Trends Microbiol* **20**: 74–79.

-
- Williams, D.R., Motallebi-Veshareh, M., and Thomas, C.M. (1993) Multifunctional repressor KorB can block transcription by preventing isomerization of RNA polymerase–promoter complexes. *Nucleic Acids Res* **21**: 1141–1148.
- Worcel, A., and Burgi, E. (1972) On the structure of the folded chromosome of *Escherichia coli*. *J Mol Biol* **71**: 127–147.
- Wu, L.J., and Errington, J. (1994) *Bacillus subtilis* SpoIIIE protein required for DNA segregation during asymmetric cell division. *Science* **264**: 572–575.
- Wu, L.J., and Errington, J. (2003) RacA and the Soj-Spo0J system combine to effect polar chromosome segregation in sporulating *Bacillus subtilis*. *Mol Microbiol* **49**: 1463–1475.
- Wu, L.J., Feucht, A., and Errington, J. (1998) Prespore-specific gene expression in *Bacillus subtilis* driven by sequestration of SpoIIIE phosphatase to the prespore side of the asymmetric septum. *Genes Dev* **12**: 1371–1380.
- Wu, L.J., Lewis, P.J., Allmansberger, R., Hauser, P.M., and Errington, J. (1995) A conjugation-like mechanism for prespore chromosome partitioning during sporulation in *Bacillus subtilis*. *Genes Dev* **9**: 1316–1326.
- Yamaichi, Y., Fogel, M.A., McLeod, S.M., Hui, M.P., and Waldor, M.K. (2007) Distinct centromere-like *parS* sites on the two chromosomes of *Vibrio* spp. *J Bacteriol* **189**: 5314–5324.
- Yamaichi, Y., Fogel, M.A., and Waldor, M.K. (2007) *par* genes and the pathology of chromosome loss in *Vibrio cholerae*. *Proc Natl Acad Sci* **104**: 630.
- Yamaichi, Y., and Niki, H. (2000) Active segregation by the *Bacillus subtilis* partitioning system in *Escherichia coli*. *Proc Natl Acad Sci U S A* **97**: 14656–14661.
- Yamaichi, Y., and Niki, H. (2004) *migS*, a cis-acting site that affects bipolar positioning of *oriC* on the *Escherichia coli* chromosome. *EMBO J* **23**: 221–233.
- Yates, P., Lane, D., and Biek, D.P. (1999) The F plasmid centromere, *sopC*, is required for full repression of the *sopAB* operon. *J Mol Biol* **290**: 627–638.
- Youngren, B., Radnedge, L., Hu, P., Garcia, E., and Austin, S. (2000) A plasmid partition system of the P1-P7*par* family from the pMT1 virulence plasmid of *Yersinia pestis*. *J Bacteriol* **182**: 3924–3928.
- Zimmerman, S.B. (2006) Shape and compaction of *Escherichia coli* nucleoids. *J Struct Biol* **156**: 255–261.